(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 December 2001 (20.12.2001)

PCT

(10) International Publication Number
WO 01/96313 A1

(51) International Patent Classification[7]: C07D 233/54, 405/14, 407/14, 403/14, 413/14, 487/02, A61K 31/478, 31/403, 31/4045, A61N 35/00

(21) International Application Number: PCT/US01/19404

(22) International Filing Date: 14 June 2001 (14.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/211,760 14 June 2000 (14.06.2000) US

(71) Applicant *(for all designated States except US)*: **THE SCRIPPS RESEARCH INSTITUTE** [US/US]; 10550 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **BOGER, Dale, L.** [US/US]; 2819 Via Posada, La Jolla, CA 92037 (US).

(74) Agents: **LEWIS, Donald, G.** et al.; The Scripps Research Institute, 10550 North Torrey Pines Road, TPC-8, La Jolla, CA 92037 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/96313 A1

(54) Title: DISTAMYCIN A ANALOGS

(57) Abstract: The development of a solution-phase synthesis of distamycin A and its extension to the preparation of 2640 analogs are described. Thus, solution-phase synthesis techniques with reaction workup and purification employing acid/base liquid-liquid extractions were used in the multistep preparation of distamycin A (8 steps, 40 % overall yield) and a prototypical library of 2640 analogs providing intermediates and final products that are ≥ 95 % pure on conventional reaction scales. Screening the prototypical library provided compounds that are 1000 times more potent than distamycin A in cytotoxic assays (67, IC_{50} = 29 nM, L1210), that bind to poly[dA]-poly[dT] with comparable affinity, and that exhibit an altered DNA binding sequence selectivity. Several candidates were identified which bound the five base-pair AT-rich site of the PSA-ARE-3 sequence, and one (128, K = 3.2 x 10^6 M^{-1}) maintained the high affinity binding (K = 4.5 x 10^6 M^{-1}) to the ARE-consensus sequence containing a GC base-pair interrupted five base-pair AT-rich site suitable for inhibition of gene transcription initiated by hormone insensitive androgen receptor dimerization and DNA binding characteristic of therapeutic resistant prostate cancer.

DISTAMYCIN A ANALOGS

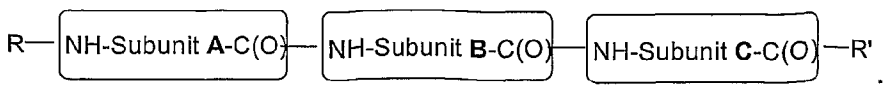
Field of Invention:

The invention relates to cytotoxic agents. More particularly, the invention relates analogs of distamycin A, to libraries of distamycin A, to their synthesis and screening for DNA binding activity and cytotoxic activies.

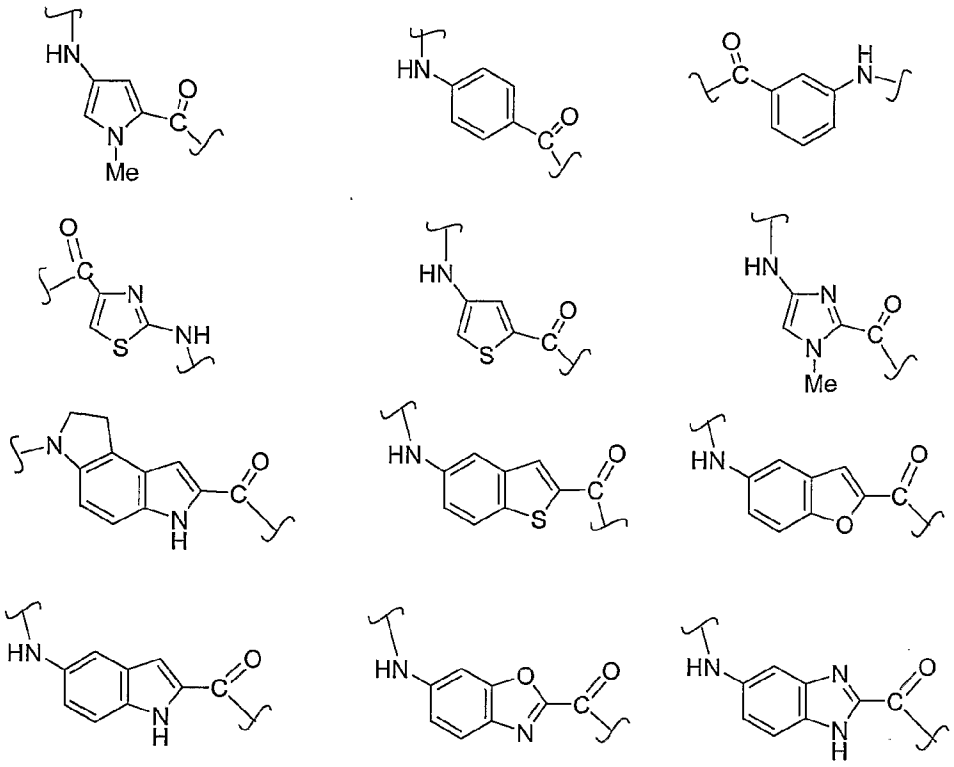
Summary:

Solution-phase combinatorial strategies for the synthesis of libraries of distamycin A analogs is described wherein the *N*-methylpyrrole subunit of distamycin A is systematically replaced with other heterocyclic amino acids. Solution-phase synthesis techniques with reaction workup and purification employing acid/base liquid–liquid extractions were used in the multistep preparation of distamycin A (8 steps, 40% overall yield) and a prototypical library of 2640 analogs providing intermediates and final products that are ≥ 95% pure on conventional reaction scales. This first generation library was further functionalized with a basic side chain to mimic the amidine group of distamycin. Screening the prototypical library provided compounds that are 1000 times more potent than distamycin A in cytotoxic assays (**67**, IC_{50} = 29 nM, L1210), that bind to poly[dA]—poly[dT] with comparable affinity, and that exhibit an altered DNA binding sequence selectivity. Several candidates were identified which bound the five base-pair AT-rich site of the PSA-ARE-3 sequence, and one (**128**, K = 3.2 x 10^6 M^{-1}) maintained the high affinity binding (K = 4.5 x 10^6 M^{-1}) to the ARE-consensus sequence containing a GC base-pair interrupted five base-pair AT-rich site suitable for inhibition of gene transcription initiated by hormone insensitive androgen receptor dimerization and DNA binding characteristic of therapeutic resistant prostate cancer.

One aspect of the invention is directed to an analog of distamycin A represented by the following structure:

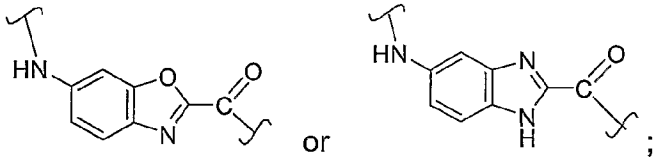


In the above structure, **R** is a selected from -C(O)O-(C1-C6 alkyl) and -C(O)$CH_2CH_2CH_2NMe_2$; **R'** is -O(C1-C6 alkyl), where (C1-C6 alkyl) is any branched or unbranched alkyl having 1 to 6 carbons; and -NH-Subunit **A**-C(O)- , -NH-Subunit **B**-C(O)- , and -NH-Subunit **C**-C(O)- are each a diradical independently selected from the following structures:

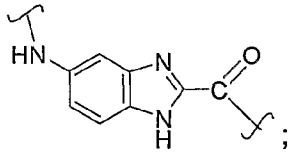


However, the following provisos apply:

1.) -NH-Subunit **A**-C(O)- can not be represented by either of the following structures:

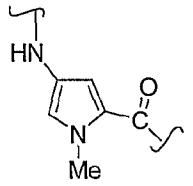


2.) -NH-Subunit **B**-C(O)- can not be represented by following structure;

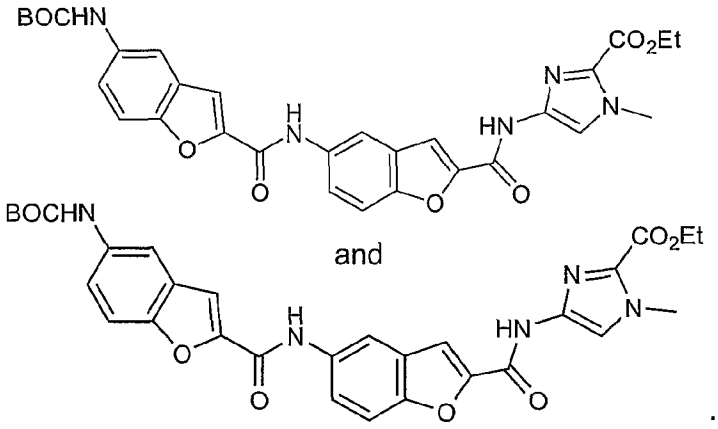


3.) -NH-Subunit **A**-C(O)- , -NH-Subunit **B**-C(O)- , and -NH-Subunit **C**-C(O)- can

not all simultaneously be represented by the following structure:

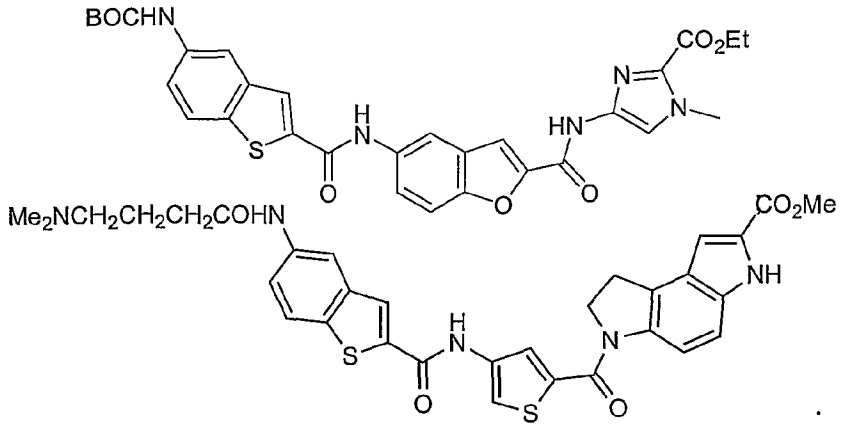


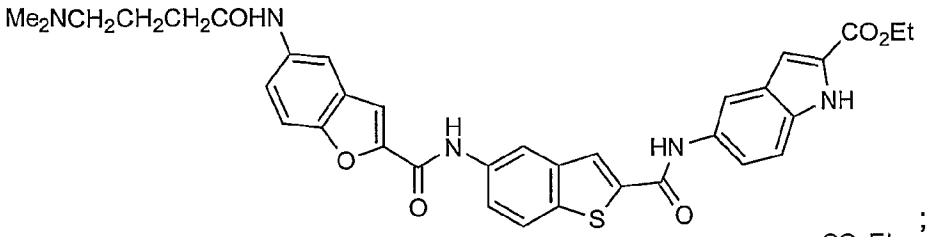
In a first preferred embodiment, **R** is -C(O)O-*t*Bu. In a second preferred embodiment, **R** is $-C(O)CH_2CH_2CH_2NMe_2$. In a third preferred embodiment **R'** is selected from the group consisting of -OMe and -OEt. In a fourth preferred embodiment, there is a further proviso that -NH-Subunit **A**-C(O)- , -NH-Subunit **B**-C(O)- , and -NH-Subunit **C**-C(O)- can not all be identical. Examples of this fourth preferred embodiment include the following species:



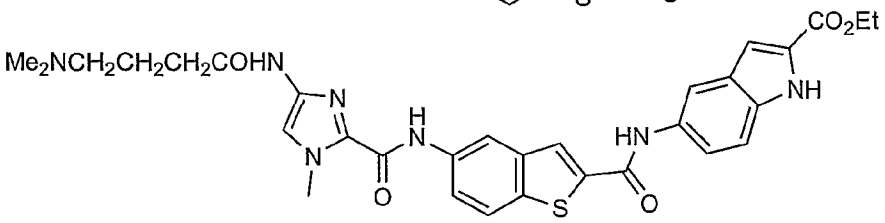
and

In a fifth preferred embodiment, there is a further proviso that none of -NH-Subunit **A**-C(O)- , -NH-Subunit **B**-C(O)- , and -NH-Subunit **C**-C(O)- are identical. Examples of this fifth preferred embodiment include the following:

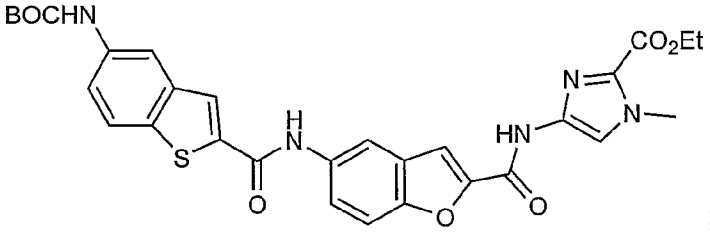




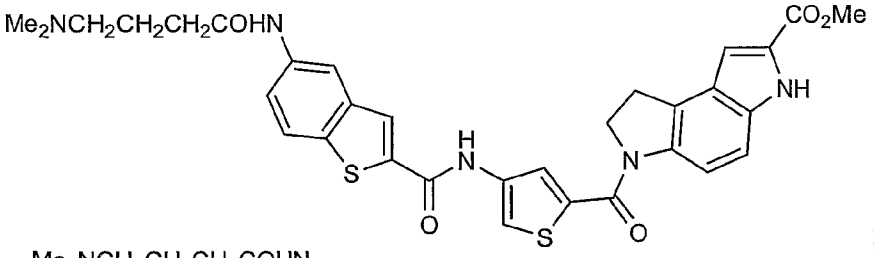
;



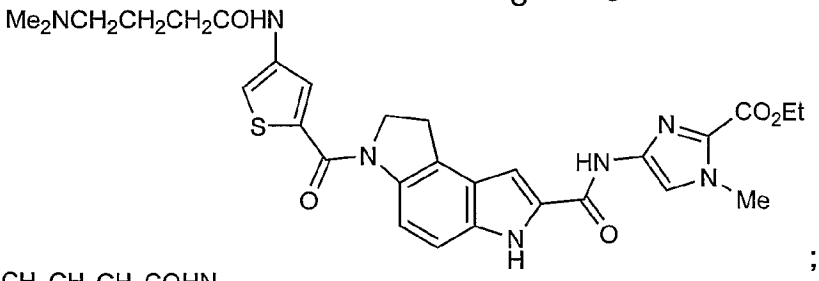
;



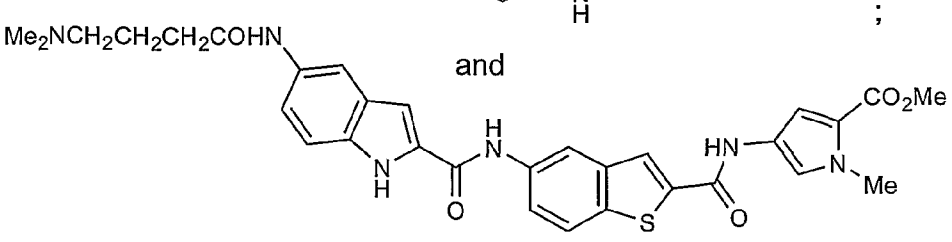
;



;



;



and

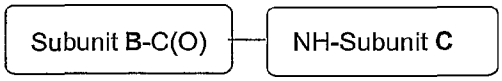
.

Another aspect of the invention is directed to a positional scanning library comprising a collection of ten or more of the compounds indicated above.

Another aspect of the invention is directed to a process for synthesizing a

library of amide linked aromatic trimers represented by the following structure:

In the above structure, Subunit **A** is any aromatic radical of a plurality of aromatic radicals, Subunit **B** is a first aromatic radical, and Subunit **C** is a second aromatic radical. In first step of the process, Subunit **B** is linked to Subunit **C** by means of a first amide linkage to form a dimer of the first and second aromatic radicals, the dimer being represented by the following structure:



In the second step of the process, a plurality of the dimers of the first step are linked to a plurality of Subunits **A** by means of a second amide linkage for forming the library of compounds. Each element of the library is a trimer of aromatic radicals linked by amide linkages.

Another aspect of the invention is directed to a process for killing a cancer cell. The process employs the step of contacting the cancer cell with a solution containing a cytotoxic concentration of any of the above compounds.

Additionally, two 1000-membered positional scanning libraries of distamycin A analogues are described and screened. The results of their screening for functional activity (L1210 cytotoxic potency) and DNA binding affinity were compared with those derived from libraries containing the same compound members but prepared in a smaller ten compound mixture format. The positional scanning libraries, which are substantially less demanding to prepare, allowed the accurate detection of the global observations and the clearly more potent activities, but more subtle discoveries and less distinguishable activities were not detected. This is a natural consequence of testing the larger 100 compound mixtures and the relative insensitivity of the assays to the contribution of any single, uniquely acting compound in the mixture. Thus, the disadvantages associated with the loss of some information contained within the library must be balanced against the advantages of the ease of library synthesis and judged in

light of the library screening objectives.

The screening of two prototypical positional scanning libraries and their comparison with prior results obtained on a library composed of mixtures of ten compounds prepared by parallel synthesis was conducted. In a cellular assay for functional activity (L-1210 cytotoxic activity), the two potent members of the library were identified but required more extensive deconvolution and were deduced from activities that exhibit less distinction. Notably, the combination of most effective residues revealed in the assay did not correspond to the most potent compound or even an effective compound. This is a natural consequence of the testing of 100 compound mixtures where the impact of any single compound is relatively small. The performance in a DNA binding assay was just as revealing. The close distamycin A analogue **17** was identified as the most effective binder to a hairpin oligonucleotide that contains the PSA-ARE-3 consensus sequence and a 5 base-pair AT-rich site. However, the distinctions in the assay were small and subtle discoveries tucked into the library were not detected, including additional effective binders and those which bound both the PSA-ARE-3 and ARE consensus sequence equally well. Nonetheless, the combined use of solution-phase mixture synthesis and positional scanning is simple and technically nondemanding even for a large library being less demanding than the parallel synthesis of individual compounds or small mixtures, or solid-phase split and mix synthesis (Furka, A.; et al. *Abstr. Int. Congress Biochem., 14th* **1988**, *5*, 47; Furka, A.; et al. *Int. J. Peptide Prot. Res.* **1991**, *37*, 487; Furka, A. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 413; Houghten, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 5131) with or without tagging (Brenner, S.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5381; Nielsen, J.; et al. *J. Am. Chem. Soc.* **1993**, *115*, 9812; Needles, M. C.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10700; Still, W. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10922; Still, W. C. *Acc. Chem. Res.* **1996**, *29*, 155). Unlike iterative (Geysen, H. M.; et al. *Mol. Immunol.* **1986**, *23*, 709; Ecker, D. J.; et al. *Nucleic Acids Res.* **1993**, *21*, 1853), surf (Wyatt, J. R.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1356), or recursive deconvolution (Erb, E.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11422), positional scanning can be conducted upfront for depository libraries subjected to

multiple screening assays. Thus, the loss of resolution in the testing data, which does not preclude identifying effective leads, must be balanced against the ease and ultimate breath of the library synthesis and its appropriateness judged in light of the screening objectives.

Brief Description of Drawings:

Figure 1 shows the design of the positional scanning library.

Figure 2 shows the structures of amino acid monomer units used in the preparation of the libraries.

Figure 3 is a scheme illustrating the synthesis of the BOC-trimer positional scanning libraries.

Figure 4 is a scheme for the conversion of BOC-trimer into DMABA-trimer libraries.

Figure 5 is a table showing the yields of the BOC- and DMABA-trimers.

Figure 6 is a bar graph showing the most potent residues that were found using the positional scanning library.

Figure 7 is a table showing the cytotoxic activities of the candidate compounds composed of the most potent residues found by the positional scanning library.

Figure 8 shows the structure of the two most cytotoxic compounds within the libraries.

Figure 9 shows the cytotoxicity (L1210) for DMABA-trimer scanning libraries. Smaller numbers indicate higher cytotoxic activity.

Figure 10 shows the structures of **86**, **210**, **220** and **49**. Compounds **86**, **210** and **220** were identified as potent DMABA-trimers.

Figure 11 shows the results of the ethidium bromide displacement assay for DMABA-trimer libraries (99 μM). DNA at 0.88×10^{-5} M, ethidium bromide at 0.44×10^{-5} M. Smaller numbers indicate higher cytotoxic activity..

Figure 12 shows the solution phase strategy for DNA binding agent libraries.

Figure 13 illustrates the general procedure for determination of sequence selectivity for a library of DNA binding agents.

Figure 14 is a scheme showing the steps in synthesizing distamycin A.

Figure 15 shows the reaction sequence for preparation of DNA binding agent libraries.

Figure 16 shows the structures of the amino acid subunits used in the preparation of libraries.

Figure 17 is a scheme for the EDCI/DMAP coupling of the carboxylic acid and amine. The second reaction was used to couple the sodium salt of carboxylic acids to amines. This method was used when the free carboxylic acids were unstable.

Figure 18 is a scheme showing the side reaction and solution to incorporating the indole subunit into analogs. The indole amino acid **13b** dimerizes upon attempted coupling with other amino acids. The series of reactions in the second and third row is how **13b** was eventually coupled to other amino acids by first protecting the indole nitrogen, coupling and then deprotecting the indole and amine nitrogens simultaneously.

Figure 19 shows the formation of the individual trimers from the pyrrole dimer. The lower reaction is the coupling of the dimers to a mixture of free acids to get mixtures of trimers.

Figure 20 is a scheme that shows the incorporation of a dimethylaminobutyric acid tail onto the deprotected trimers.

Figure 21 is a three-dimensional display of the results of the cytotoxicity assay for the BOC-trimer libraries.

Figure 22 shows how the individual trimers were synthesized after finding out which were the best B and C subunits. The activity of each individual trimer is shown in the table.

Figure 23 is a three-dimensional display of the results of the cytotoxicity assay of the dimethylaminobutyric acid terminated trimers. A table shows the toxicity of the individual compounds.

Figure 24 shows the synthesis of the individual dimethylaminobutyric acid (DMABA) terminated trimers for testing. The most active B and C subunit was determined in the cytotoxicity assay.

Figure 25 shows the general procedure for establishing DNA binding of a

library of compounds with a single sequence.

Figure 26 shows the ethidium bromide displacement assay for DMABA-trimer libraries with poly[dA]-poly[dT] DNA in graph **A** and graph **B** shows the corresponding assay for poly[dG]-poly[dC]. Larger numbers indicate higher affinity for DNA.

Figure 27 shows the synthesis of 40 DMABA trimers and the corresponding yields.

Figure 28 shows the activity in the L1210 assay and the percent remaining fluorescence with poly[dA]-poly[dT] DNA.

Figure 29 shows the graphical results for the ethidium displacement assay for selected DMABA trimers. The hairpin oligonucleotides contain the 14-base pair ARE-consensus and the PSA-ARE-3 sequences.

Figure 30 shows the results of the ethidium bromide assay for selected compounds in table form.

Figure 31 gives the hairpin structure of the nucleotides representing all possible combinations of five base pairs.

Figure 32 shows the graphical results of a screen of distamycin A against the library of DNA hairpin oligonucleotides.

Figure 33 is a table with the binding constants of distamycin A with particular short AT-rich sequences.

Figure 34 is a table with binding constants of different types of DNA with ethidium bromide.

Figure 35 shows the results of a screen of compound **128** with a library of 512 DNA hairpin oligonucleotides. The top 20 sequences are shown.

Figure 36 shows the binding constants with two types of DNA, Gibb's free energy of binding to poly[dA]-poly[dT] DNA, and the IC_{50}'s from the L1210 assay of a selected group of 6 compounds which are analogs of high affinity DNA binding agents.

Figure 37 shows the binding constants with two types of DNA, Gibb's free energy of binding to poly[dA]-poly[dT] DNA, and the IC_{50}'s from the L1210 assay of another group of 6 compounds which are analogs of high affinity DNA binding agents.

Detailed Description:

A procedure for rapidly determining DNA binding selectivity entailing the competitive displacement of prebound ethidium bromide from defined hairpin oligonucleotides (Figure 13). (*Drug–DNA Interactions Protocols*; Fox, K. R., Ed.; Methods in Molecular Biology; Humana Press: Totowa, New Jersey, 1997; Vol. 90; Jenkins, T. C. Optical Absorbance and Fluorescence Techniques for Measuring DNA-Drug Interactions, In *Drug–DNA Interactions Protocols*; Fox, K. R. Ed.; Methods in Molecular Biology; Humana Press: Totowa, New Jersey, 1997; Vol. 90, p 195.; Morgan, A. R., et al., *Nucleic Acids Res.* **1979**, *7*, 547; Baguley, B. C., et al., *Nucleic Acids Res.* **1978**, *5*, 161; Boger, D. L., et al., *Chem.-Biol. Interact.* **1990,** *73*, 29; and Boger, D. L., et al., *J. Org. Chem.* **1992**, *57*, 1277). DNA of interest (homopolymers, heteropolymers, or predefined hairpin oligonucleotides) in 96-well plates is treated with ethidium bromide, yielding a large fluorescence increase upon DNA intercalation. Addition of a nonfluorescent DNA binding agent results in a decrease in fluorescence due to displacement of bound ethidium bromide. The decrease in % fluorescence is directly related to the extent of DNA binding providing relative DNA binding affinities and, through subsequent quantitative titration, is capable of providing accurate absolute binding constants.

As detailed herein, this technique may be used to screen a library of compounds for DNA binding to a single DNA sequence or for the complementary screening a single compound against a full library of DNA sequences which results in the definition of the sequence specificity of a given agent. Combining these in the assay of a library of compounds against a library of DNA, provides qualitative and/or quantitative information on the binding of all library members against a library of available sequences, allowing complete characterization of the DNA binding profiles of each agent in a single experiment.

Solution-phase Total Synthesis of Distamycin A.

As an initial demonstration of the approach, we first conducted a total synthesis of distamycin A utilizing solution-phase synthesis techniques that require only acid/base liquid–liquid extraction purification protocols. Previous total

syntheses of distamycin A generally have relied on the coupling of *N*-methyl-4-nitropyrrole-2-carboxylic acid chlorides followed by nitro group reduction and further coupling steps (Lown, J. W., et al., *J. Org. Chem.* **1985**, *50*, 3774; Grehn, L., et al., *J. Org. Chem.* **1981**, *46*, 3492; and Bialer, M., et al., *Tetrahedron* **1978**, *34*, 2389). In developing a general set of reaction conditions suitable for the preparation of libraries, several requirements not intrinsic to the natural product synthesis needed to be addressed. In all cases, unreacted starting materials, coupling agents, and their reaction by-products needed to be removed by simple acid/base extraction. Although acid chlorides could be used, the preparation and long-term storage of numerous heterocyclic acid chlorides would be difficult, requiring the implementation of coupling protocols that use the carboxylic acids directly. In addition, a nitro group reduction step introduces a reaction of variable generality (reaction time, catalyst poisoning), precludes the inclusion of subunits sensitive to reduction conditions, and requires the resultant free amines to be relatively stable. Consequently, we adopted the more direct use of BOC protected amines. A general set of coupling conditions that gives high yields of coupled product was developed enlisting the amines and carboxylic acids directly and the water soluble 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride (EDCI) with dimethylaminopyridine (DMAP) as an additive. The unreacted starting materials, reagents, reaction and reagent by-products all may be removed by acid/base extractions.

Starting with the pyrrole carboxylic acid **1a**, coupling with aminopyrrole **1b** using EDCI/DMAP afforded **2** in high yield (97%). Removal of the BOC protecting group with HCl/EtOAc followed by coupling to pyrrole **1a** afforded the tripeptide **3** in good yield (96%). Saponification of **3** followed by coupling with β-aminopropionitrile afforded nitrile **4** in excellent yield (95%). Treatment of nitrile **4** with HCl/EtOH followed by NH_3/EtOH afforded the desired amidine with concomitant removal of the BOC group. Due to the intrinsic instability of this free amine, it was immediately treated with *N*-formyl imidazole to afford distamycin A. This provided distamycin A in 40% overall yield for eight steps without deliberate optimization, and required only acid/base liquid–liquid extraction to afford all intermediates and the final product with >95% purity as demonstrated by their ^1H

NMR spectra.

Library Design.

Two prototypical libraries of potential DNA binding agents were prepared in a small mixture format, Figure 15. Using eleven *N*-BOC heterocyclic amino acids and twelve amino esters, the individual subunits were coupled using EDCI/DMAP to provide all possible 132 individual dipeptides in parallel. The use of EDCI and DMAP allows for the removal of excess coupling agents and their reaction by-products along with unreacted starting materials by acid/base liquid–liquid extraction. These individual dimers were deprotected and coupled to a mixture of ten *N*-BOC carboxylic acids to give 132 mixtures of ten *N*-BOC-trimers where only the last position (subunit A) is undefined (1320 compounds). Removal of the BOC group and coupling to the basic side chain, *N, N*-dimethylaminobutyric acid (DMABA), affords an analogous DMABA-trimer library (1320 compounds). The amidine group found in distamycin A was replaced with a dimethylamino group in order to avoid the variable yielding Pinner reaction and for overall ease of synthesis. The decision to place the basic side chain at the *N*-terminus rather than the *C*-terminus resulted from observation of inefficiencies during hydrolysis of certain *C*-terminal monomer subunits (see below). Comparison studies detailed in a following section determined that these changes had little effect on the DNA binding affinities of the resultant agents.

This strategy of preparing all individual dimers offered the advantage of examining the reactivity of all possible combinations of acids and amines to ensure that the coupling conditions and purification protocols were general and provided the levels of purity desired. Since the reactions are carried out in solution, each of 132 dimers were characterized by ^{1}H NMR (see Supporting Information), purity was established by conventional techniques, and 50–100 mg quantities were accessible. This allows for the preparation of numerous second generation libraries from stocks of dimers and for the deconvolution of libraries from stored and archived samples of all library components.

The heterocyclic amino acids selected for the first prototypical libraries are shown in Figure 16. Included in this set are the pyrrole, imidazole, and thiazole

amino acids studied by Dervan and Lown, and the indole and CDPI amino acids studied in our laboratories. While not intended to be a survey of optimal heterocyclic amino acids, this set provides built-in known DNA binding agents and could be expected to address issues of identification, practicality and viability which proved to be especially useful in comparisons with positional scanning libraries.

The subunits **1**, **5**, **6**, **9**, **10**, and **13** were prepared according to known procedures (Baird, E. E., et al., *J. Am. Chem. Soc.* **1996**, *118*, 6141; Nishiwaki, E., et al., *Heterocycles* **1988**, *27*, 1945; Boger, D. L., et al., *J. Org. Chem.* **1987**, *52*, 1521; and Boger, D. L., et al., *Bioorg. Med. Chem.* **1995**, *3*, 1429). The preparation of the remaining subunits is obtained from readily available materials following established procedures and proceed through intermediates **16–29** (Sprague, J. M., et al., *J. Am. Chem. Soc.* **1946**, *68*, 266; Foye, W. O., et al., *J. Am. Chem. Soc.* **1954**, *76*, 1378; and Shih, C., et al., *J. Med. Chem.* **1992**, *35*, 1109; Osuga, H., et al., *Bull. Chem. Soc. Jpn.* **1997**, *70*, 891; Van Wijngaarden, I., et al., *J. Med. Chem.* **1988**, *31*, 1934; Moller, H. *Liebigs Ann. Chem.* **1971**, *749*, 1; Crivello, J. V. *J. Org. Chem.* **1981**, *46*, 3056; Bistrzycki, A., et al., *Chem. Ber.* **1912**, *45*, 3483; Boger, D. L., et al., *Bioorg. Med. Chem.* **1995**, *3*, 761; Rastogi, R., et al., *Ind. J. Chem., Sect. B* **1979**, 464). Ester hydrolysis of the precursors to both **14** and **15** proceed efficiently with NaOH in MeOH, but attempts to isolate the free carboxylic acids led to decarboxylation. However, the sodium salts **14b** and **15b** were isolated in quantitative yield without detectable decarboxylation and used effectively in our efforts.

Parallel Synthesis of Dimers.

Using **1b** and **5b–13b** as the acid component and **1a** and **5a–15a** as the amine component, 120 individual dimers were prepared (Figure 17). Each dimer was prepared in 70–80 mg quantities in parallel, using only acid/base liquid–liquid extraction purification to afford products in typically >80% yield and >95% purity.

Incorporating the benzoxazole **14** into the dimers presented a unique problem as attempts to isolate the free acid (**14**, R' = H) led to decarboxylation. It was possible to isolate the sodium salt (**14b**, R = Na), but the reaction conditions

used to prepare the individual dimers in Figure 17 (EDCI, DMAP) led to complex reaction mixtures and significant amounts of the decarboxylated benzoxazole. However, PyBrop was found to work well in their preparation and was used in lieu of the standard conditions (Figure 17). In three instances, when using the indole subunit **13b** as the acid component in couplings with the three unreactive amines (**5a**, **7a**, and **14a**), the diketopiperazine **30** was isolated due to indole dimerization (Figure 18). Although not the topic of the present work, the properties of the indole diketopiperazines have proved extraordinary, providing potent cytotoxic agents displaying effective DNA binding properties in their own right (Boger, D. L., et al., *Bioorg. Med. Chem. Lett.* **2000**, *10,* 0000). To circumvent this problem, the indole nitrogen was protected with a *p*-methoxybenzyl group to afford indole **31**. Hydrolysis to afford the free acid **32** and coupling to the three individual amines afforded the desired dipeptides in moderate yield. Simultaneous deprotection of both the *p*-methoxybenzyl and BOC-protecting groups (TFA/anisole, 60 °C) afforded the desired amines. We were unable to prepare dimers where benzimidazole **15b** was the acid component. Therefore, this monomer was only used in the third position (C) of the trimers.

Synthesis of BOC-trimer Libraries. The preparation of the trimer libraries was investigated initially by preparing several sets of individual trimers to ensure the reaction conditions were appropriate. For example, the preparation of all ten individual trimers from the BOCNH-**1**-CONH-**1**-OMe dimer is given in Figure 19. Deprotection of the dimer with HCl/EtOAc followed by coupling with **1b**, **5b–13b** afforded the ten BOC-trimers in high yield and with >90% purity using only acid/base liquid–liquid extraction purification. This set based on the dipyrrole dimer is of special interest because it contains a close analog of distamycin, the tripyrrole **39** (BOCNH-**1**-CONH-**1**-CONH-1-OMe).

Having established the conditions for coupling and workup, each of the individual dipeptides was converted to a mixture of ten tripeptides (Figure 19). Removal of the BOC group with anhydrous HCl, followed by coupling to an equimolar mixture of ten acids afforded the BOC-trimer mixtures on a 50 μmole scale. An excess of the amine component was used to ensure complete consumption of the ten acids in the reaction mixture. Full matrix mixtures

analyzed by mass spectrometry ensured all expected components were present (see Supporting Information). Since the benzoxazole **14b** did not couple efficiently under the standard conditions and required an additional purification step, it was omitted from the first position (subunit A) to ensure library fidelity and purity.

Dimethylaminobutyric Acid Libraries. In a similar manner, optimization of reaction conditions to incorporate a dimethylaminobutyric acid side chain (DMABA) was carried out on individual trimers (Figure 20). The yields were found to vary more widely than that of the previous steps since some of the derivatives showed appreciable water solubility. Thus, the typical acid/base liquid–liquid purification protocol was modified. Simply removing the solvent from the reactions, followed by suspension of the products in water and extraction with EtOAc gave the desired products.

Using the conditions described for the individual compounds, each of the BOC-trimer mixtures was converted to the corresponding DMABA-trimer mixture as shown in Figure 20. Full matrix mixtures analyzed by mass spectrometry ensured all components were present (see Supporting Information). With the DMABA-trimer library, 2640 distamycin analogs were available in the format of two small mixture libraries.

Cytotoxic Activity.

In addition to our interest in the DNA binding properties of the compounds, we were also interested in comparing the behavior of the libraries in functional assays. In particular, we were interested in comparing the performance of the library format of 10 compound mixtures versus larger mixture testing required of positional scanning. Consequently, the two libraries were examined in a functional assay for cytotoxic activity (L1210) (Boger, D. L., et al., *J. Med. Chem.* **1985**, *28*, 1543). While all of the BOC-trimer mixtures showed some activity in the cell-based assay (IC_{50} < 10 μM), thirteen showed activity at less than 1 μM, and one showed activity at less than 100 nM (Figure 21). The most active library contained the benzofuran subunit (**12**) at the central position (subunit B) and the imidazole subunit (**9**) at the final position (subunit C).

This mixture was deconvoluted by resynthesis of the ten components,

beginning with the stored BOCNH-**12**-CONH-**9**-OEt dimer (Figure 22). This resynthesis from the immediate precursor required a single day and provided 5 mg samples of each component. A second round of testing revealed that the most active components contained either the benzofuran (**12**) or the benzothiophene (**11**) at the first (A) position, with IC_{50}'s of 29 nM and 68 nM for **66** and **67**, respectively (Figure 22). When compared to distamycin A (IC_{50} = 42 μM), both are 1000 times more potent. For comparison purposes, **39–48**, which contain the dipyrrole subunits of distamycin and its close tripyrrole analog **39**, were also examined in the L1210 assay (Figure 21). Consistent with the behavior of distamycin A, **39** was essentially inactive (IC_{50} = 32 μM) and approximately 1000 times less potent than **66** and **67**.

In contrast, the IC_{50} values for the DMABA-trimer mixtures were found to be on the order of 10–100 fold higher than the BOC-trimer libraries (Figure 23). This may be the result of decreased cell penetration of a charged species. The most active mixture contains the CDPI subunit (**10**) in the final position (subunit C), and the thiophene subunit (**8**) at the central position (subunit B). This mixture was deconvoluted by synthesis of the ten components, beginning from the BOCNH-**8**-CONH-**10**-OMe dimer (Figure 24). A second round of screening revealed that the most active component of this library contained the benzothiophene (**11**) at the first position (subunit A), with an IC_{50} of 0.46 μM for **86** and it was >10 times more active than any other compound in the mixture and 100 times more potent than any of the individual components of the **49–58** mixture based on and including the close distamycin analogs (Figure 23). The tripyrrole analog **49** exhibited an IC_{50} of 44 μM indistinguishable from that of distamyin A (42 μM) and 100 times as less potent than **86**.

In the instances examined, including additional mixtures that were deconvoluted while examining the DNA binding properties of the agents (see Figure 27), the activity of the mixtures in the cell-based assay approximated that of the individual components and established the reliability of testing in the small mixture format for the libraries.

The individual BOC-dimers were also tested in the L1210 functional assay. Nearly all the members were essentially inactive (IC_{50} > 1μM) with the exception

of those containing the B subunit **6** and the bicyclic heterocycles **10–15** in the final position (subunit C), five of which exhibited IC_{50}'s < 1μM. The most potent, BOCNH-**6**-CONH-**10**-OMe, exhibited superb activity with an IC_{50} = 28 nM being >1000 times more active than distamycin or its dipyrrole analog BOCNH-**1**-CONH-**1**-OMe.

DNA Binding Studies. The most interesting opportunities for us lies with the establishment of the DNA binding properties of the library members. Although a variety of techniques are commonly used to investigate the DNA binding properties of small molecules, most are technically challenging and time consuming, making them inapplicable to high-throughput screening. However, one technique entailing the competitive displacement of prebound ethidium bromide, does represent a potentially useful high-throughput assay when used in conjunction with a 96-well fluorescent plate reader (Figure 13). Ethidium bromide yields a large fluorescence increase upon DNA intercalation. Addition of another nonfluorescent DNA binding agent results in a decrease in fluorescence due to displacement of bound ethidium bromide. The procedure provides a rapid, flexible, and reliable indication of the relative binding affinities of a wide variety of DNA binding ligands. This technique was first examined as a rapid screen for binding to poly[dA]–poly[dT] and poly[dG]–poly[dC] and subsequently extended to hairpin oligonucleotides containing unique sequences. Using a single agent concentration, the relative decrease in % fluorescence is proportional to the affinity of a given mixture or individual compound for a particular DNA sequence. In addition, we have found that the well-defined linear reduction in fluorescence upon titration with agents related to distamycin can be used to establish absolute binding constants. This ability to provide both relative and absolute binding constants enlisting a technically nondemanding assay provides a powerful screening complement to the library synthesis.

Poly[dA]–Poly[dT] and Poly[dG]-Poly[dC].

The binding results for the DMABA-trimer library with poly[dA]–poly[dT] showed several general trends: (1) all the DMABA-trimers induce some decrease in fluorescence, indicating the libraries have an overall AT affinity; (2) high affinity

libraries contain one of the larger subunits at the second (B) position (monomers **10–14**); and (3) the smaller subunits at the third (C) position (monomers **1, 5–9**) appear to be more active. Notably, four of the 10 compound mixtures showed a higher affinity than the pyrrole sublibrary containing **49**, the tripyrrole analog of distamycin. The highest affinity mixture contains the CDPI subunit (**10**) at the second position and the imidazole subunit (**9**) at the third position. The second most effective mixture contains the benzothiophene (**11**) at the second position and the pyrrole (**1**) at the third position.

Both were deconvoluted through parallel synthesis of the ten individual components (Figure 27) and their assay revealed that **112** and **128** showed the highest affinity for poly[dA]–poly[dT]. Quantitative ethidium bromide titration of **112** and **128** afforded binding constants of $2.5 \times 10^6\ M^{-1}$ and $5.6 \times 10^6\ M^{-1}$, respectively. The latter agent proved essentially indistinguishable from the tripyrrole analog of distamycin A (**49**, $K = 5.9 \times 10^6\ M^{-1}$). Similarly, **79–88** which constitute the individual compounds of the mixture that exhibited the most potent cytotoxic activity of the DMABA-trimers were also examined and the results are recorded herein. In those instances where the monomer **10** was present in the mixture, longer incubation times were required due to insolubility. Consistent with its cytotoxic activity, **86** exhibited effective binding to poly[dA]–poly[dT].

The DMABA-trimers were also screened for binding to poly[dG]–poly[dC]. As expected, the affinity is much lower than for poly[dA]–poly[dT]. The BOC-trimer libraries were also screened for DNA binding to both poly[dA]–poly[dT] and poly[dG]–poly[dC] (data not shown), and they showed substantially lower affinity as expected.

Defined Sequence Within a Hairpin Oligonucleotide.

Although general trends may be detected by examining the binding characteristics with homopolymer DNAs, the most useful information is derived by examining their binding at defined sequences. The extension of the rapid screening to such individual sequences is illustrated with two hairpin oligonucleotides containing two related sequences of the androgen response element, the 14 base pair ARE-consensus and PSA-ARE-3 sequences (Cato, A.

C. B., et al., *EMBO J.* **1987**, *33*, 545; and Cleutjens, K. B. J. M., et al., *Mol. Endocrinology* **1993**, *7*, 23). The emergence of hormone independent, constituently active androgen receptor dimer, unresponsive to competitive antagonist treatment, is responsible for prostate cancer relapse resistant to chemotherapeutic treatment (Chang, C. C., et al., *Crit. Rev. Eucaryotic Gene Expression* **1995**, *5*, 97; Bentel, J. M., et al., *Endocrinology* **1996**, *151*, 1; Veldscholte, J., et al., *Biochem. Biophys. Res. Commun.* **1990**, *173*, 534; and Galbraigth, S. M., et al., *Eur. J. Cancer* **1997**, *33*, 545). A potentially effective treatment for such resistant prostate cancer could entail administration of a DNA binding agent selective for both the PSA-ARE-3 and ARE-consensus sequences that would competitively inhibit the androgen receptor DNA binding and its transcription activation.

Screening the entire DMABA-trimer library for binding to the two hairpin oligonucleotides, enlisting the ethidium bromide displacement assay, revealed that the mixture containing the pyrrole subunit (**1**) at both the second (B) and third (C) position gave the largest decrease in fluorescence with the PSA-ARE-3 hairpin oligonucleotide, which contains a 5 base-pair AT-rich site. Substitution of a single AT base pair with a GC base pair (ARE-consensus) at the center of this sequence results in loss of affinity for this mixture. Screening the individual components of this mixture afforded the direct distamycin A analog **49** as having the highest affinity, followed closely by **53** containing the thiophene subunit at the first position (A). The same overall pattern was observed with poly[dA]–poly[dT], where **49** and **53** also showed the highest affinity (Figure 27).

Both these agents exhibited diminished affinity for the ARE-consensus sequence presumably resulting from the intervening GC base pair. Two additional mixtures, 109–118 and **119–128**, also bound the PSA-ARE-3 sequence effectively with the general trend **119–128** > **49–58** > **109–118**, the same general trend seen with poly[dA]–poly[dT] (Figure 27). The individual trimers **124** and **128** displayed tight binding to the PSA-ARE-3 sequence analogous to **49** and **53**. Importantly, **124** showed a loss of affinity to the ARE-consensus analogous to **49** and **53**, but **128** retained equal affinity making this agent ideal in maintaining high affinity for both the PSA-ARE-3 and ARE-consensus sequences.

This raised the issue of the DNA binding selectivity of **128** and its distinctions from **49** or distamycin A that are responsible for the PSA-ARE-3 and ARE-consensus binding. Consequently, the ethidium bromide displacement assay was enlisted to define the complete sequence selectivity of both distamycin A and **128**.

Complete Sequence Selectivity of a Prototypical DNA Binding Agent Distamycin A: Rank Order Binding to a Library of Hairpin Oligonucleotides. Distamycin A is among the best characterized DNA binding compounds. Its DNA binding properties have been studied in depth through footprinting (Portugal, J., et al., *Eur. J. Biochem.* **1987**, *167*, 281; Portugal, J., et al., *FEBS Lett.* **1987**, *225*, 195; Abu-Daya, A., et al., *Nucleic Acids Res.* **1995**, *23*, 3385; Abu-Daya, A., et al., *Nucleic Acids Res.* **1997**, *25*, 4962), calorimetry (Rentzeperis, D., et al., *Biochemistry* **1995**, *34*, 2937), NMR (Pelton, J. G., et al., *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5723; Klevit, R. E., et al., *Biochemistry* **1986**, *25*, 3296; Pelton, J. G., et al., *J. Am. Chem. Soc.* **1990**, *112*, 1393), and X-ray crystallography (Coll, M., et al., *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8385). However, even for distamycin A, a detailed study of its rank order binding to all possible of DNA sequences has not been described and its affinity for nonoptimal binding sites is not easily assessed using common techniques. Consequently, it represents an ideal example with which the ethidium bromide technique could be examined in efforts to assess its use for establishing DNA binding selectivity. Thus, a survey of distamycin A binding to all possible 5 base pair DNA sequences was conducted using a library of 512 hairpin DNA oligonucleotides containing all possible five base pair sequences of the general format 5'-GCXXXXXC-3' with a 5-A loop. Although there are 1024 possible sequences containing 5 base pairs, two complementary sequences are contained in each hairpin differing only in their location relative to the position of the adenine loop making, for example, the sequence 5'-ATGCA equivalent to the sequence 5'-TGCAT.

The results of screening the 512-membered library of hairpin oligonucleotides using distamycin A is given in Figure 32. As expected, affinity increases with increasing AT content. The top sequences include the sites 5'-ATAA, 5'-AATT, 5'-AAAT, and 5'-AAAA and among the twenty hairpins showing

the greatest decrease in % fluorescence, three four base-pair sequences occur most often: 5'-AATT, 5'-AAAT, 5'-AATA.

Although distamycin A has been studied in extensive detail, suprisingly few absolute binding constants for binding to short AT-rich sequences have been published. The comparison of all those disclosed show the relative trend 5'-AATTT>AAAAA>AATAA>ATTAA (Figure 33) (Wade, W. S., et al., *Biochemistry* **1993**, *32*, 11385). The ethidium bromide displacement assay revealed the same general trend and a quantitative titration measurement of binding constants with the hairpin oligonucleotides containing these sequences afforded binding constants that are not only consistent with the relative trend (Figure 32), but also within a factor of 2–3 of *all* the absolute binding constants previously determined through calorimetry and footprinting (Figure 33). Given that the DNA upon which the measurements were made is different, that the buffer conditions are not identical, and that entries 2–4 in Figure 33 were derived from a close analog of distamycin A, all which may contribute to small discrepancies in the absolute binding constants, the ethidium bromide displacement titration assay appears to be remarkably accurate at reproducing absolute binding constants.

Because the fluorescence derived from ethidium bromide binding varies from sequence to sequence, it is the % fluorescence decrease and not the final absolute fluorescence that is proportional to the extent of DNA binding. For tight binding agents like distamycin A which display a well defined linear loss of fluorescence, the absolute binding constants can be established by quantitative titration independent of a consideration of the ethidium bromide binding constants using a noncompetitive model of K = 1/[agent]–0.5[DNA]r where K = binding constant, [agent] = concentration at 50% reduction in fluorescence, [DNA] = DNA concentration, and r = 1/binding site size (Boger, D. L., et al., *Chem.-Biol. Interact.* **1990,** *73*, 29; and Boger, D. L., et al., *J. Org. Chem.* **1992**, *57*, 1277). The latter can be determined experimentally from the extrapolated x-intercept (0% fluorescence) of a % fluorescence vs [agent]/[base pair] plot. In the present study, this was established to be 0.125 which corresponds to 1 bound agent/8 base pairs or 1 agent bound per hairpin oligonucleotide. The alternative use of a competitive binding model to calculate absolute binding constants requires a

knowledge of the ethidium bromide binding constants for each sequence and follows from $K = K_{EB}[EB]/[agent]$ where K = binding constant, K_{EB} = binding constant for ethidium bromide, [EB] = ethidium bromide concentration, [agent] = agent concentration at 50% fluorescence. The ethidium bromide binding constant varies considerably (Figure 34) and the displacement does not follow a 1:1 stoichiometry, both of which complicate the use of a competitive binding model for establishing binding constants. The most accurate usage likely would employ the calf thymus K_{app} (10×10^6 M^{-1}) which represents an average binding constant. Neglecting the stoichiometry of the displacement as recommended, this provides a $K = 7.9 \times 10^7$ M^{-1} for distamycin A binding to AATTT, essentially indistinguishable from the value established using the noncompetitive model (9.4×10^7 M^{-1}). The advantage of this method is that it does not require establishment of binding site size for the compound, but this necessarily introduces inaccuracies due to this assumption. Thus, while the competitive model provides reasonable constants for high affinity sequences, it overestimates the binding constants for the weaker sequences. Therefore, we prefer and recommend the use of the noncompetitive binding model for establishing absolute binding constants with hairpin oligonucleotides and caution that they should be further validated by other techniques.

Notably, there is more information, or rather a higher resolution of information, regarding the selectivity of distamycins binding to DNA in this single experiment than may be found in all past work combined, which typically is limited to identification of the highest affinity sites. The establishment of rank order affinity for all possible binding sites including the ability to determine relative *or* absolute binding constants for even modest or low affinity sites provide a new opportunity to qualitatively or quantitatively compare the DNA binding selectivity of various compounds. For example, simply the shape of the merged bar graph shown in Figure 32, the curve and slope of the resulting graph, and the area under the curve provide means by which to qualitatively and quantitatively compare selectivities of DNA binding. These are presently under investigation and our assessments of their value will be disclosed in due course.

Sequence Selectivity Determination for 128: A Novel DNA Binding Agent.

As detailed in an earlier section, **128** bound poly[dA]–poly[dT] with an affinity equal to the distamycin analog **49.** However, it also bound poly[dG]–poly[dC] with only a slightly reduced affinity being 25–30 times more effective than **49** and, unlike **49**, it bound tightly to both the PSA-ARE-3 and ARE consensus sequences. The sequence selectivity of **128** was established by screening it against the library of 512 hairpin oligonucleotides. Compound **128** was found to clearly bind with a selectivity distinct from that of distamycin A and it appears to exhibit a significant preference for PuPyPy sequences. Of the 20 highest affinity sequences, sixteen contain the PuPyPy motif (80%), where statistically 37.5% of the sequences would be expected to contain this motif in a random sample. One of the four exceptions contained a 5 base pair AT rich site. Within both of the androgen response elements used to identify **128**, the PuPyPy motif is repeated three times. It appears this may be the reason for the equally high affinity binding of **128** with both sequences. Further studies on **128** are required to establish the structural basis for its DNA binding selectivity and these will be disclosed in due course.

Analysis of Side Chain Position and Nature of the Basic Functionality. Finally, in order to evaluate the effects of the structural changes that were made relative to distamycin A in the nature and position of the charged functionality, a number of derivatives of the high affinity DNA binding agents including the tripyrrole core of distamycin were prepared (Figures 36 and 37).

Analysis of these derivatives using the quantitative titration with displacement of prebound ethidium bromide showed that there is very little difference between an amidine as the basic side chain and the dimethylamino group (**Distamycin** vs. **130**) or between placing the basic side chain at the *C*- or *N*-terminal end of the trimer (**49** vs. **129**), Figure 36. Amine **129** shows slightly lower binding affinity to poly[dA]–poly[dT] than **49** which may arise from the incorporation of a bulky *t*-butyl group present at the *N*-terminus. Interestingly, there is approximately a three-fold difference in binding between distamycin A or **130** and the tripyrrole **49**. The primary difference between the two molecules is the presence of an additional potential hydrogen bond donor in distamycin and

130 (*N*-terminal formamide) that is not present in either **49** (*C*-terminal ester), **129** (*N*-terminal BOC-group), or **132** (*C*-terminal dimethylamide). When a potential hydrogen bond donor group is included at the *C*-terminus (**131**), the binding affinity does approximate that of distamycin A. The difference in free energy of binding between those molecules containing an additional donor hydrogen bonding group (distamycin A and **130**–**131**) and those which do not (**49**, **129**, **132**) is approximately 1 kcal/mol, the value of a single hydrogen bond. Interestingly, adding a second substituent containing an additional basic, protonated amine (**133** vs **131**) does not further increase the DNA binding affinity.

Analogous observations were made with derivatives of **128** which bound to poly[dA]–poly[dT] within a factor of 2 of the corresponding distamycin A derivatives. The distinctions between **128** and the distamycin derivatives were that the former bound poly[dG]–poly[dC] 15–30 times more effectively.

Conclusions.

An approach to the rapid, parallel solution-phase synthesis of distamycin A analogs was developed enlisting a simple, acid/base liquid–liquid extraction for purification and isolation of each intermediate and final product (≥ 95% pure). Its utility was demonstrated with the preparation of distamycin A and a prototypical library of 2640 analogs assembled in a small mixture format of two libraries of 132 mixtures of 10 compounds providing each in multimilligram quantities sufficient for screening in multiple assays. Screening of the library in a functional assay for cytotoxic activity (L1210) revealed two uniquely active compounds, **66** and **67**, which were 1000 times more potent than distamycin A, and **86**, which was 100 time more potent than distamycin A. More fundamental, a complementary rapid, high-throughput screen for DNA binding affinity was developed based on the loss of fluorescence derived from displacement of prebound ethidium bromide which is applicable for assessing binding to DNA homopolymers or specific sequences (hairpin oligonucleotides). Using this technique, the distamycin A tripyrrole analog **49** as well as alternative AT-rich binding agents were identified (**112** and **128**) establishing the validity of the technique and providing two new and effective DNA binding agents. In addition, a comparison of several distamycin analogs

established substituent contributions to AT-rich binding that may be safely implemented in future libraries. Extension of these studies to identify effective binders to predefined sequences was conducted in the context of the androgen response elements PSA-ARE-3 and the ARE-consensus sequences, the latter of which contains a GC base pair interrupted AT-rich sequence, and for which effective binders might prove useful in the treatment of hormone antagonist resistant prostate cancer. Three agents, **124**, **128**, and the distamycin analog **49** were identified as high affinity binders for the PSA-ARE-3 sequence incorporating a five base pair AT-rich sequence. Fundamentally more important and unlike the distamycin analog **49**, **128** retained this high affinity binding to the ARE-consensus sequence, which contains a GC base pair interrupted AT-rich site, suggesting it may serve as an effective inhibitor of androgen receptor DNA binding and its initiated gene transcription. Extension of the DNA binding assay to a powerful technique for establishing the DNA binding selectivity of an individual compound was developed enlisting distamycin A and its comparison with **128** requiring the assay of each compound against a library of 512 hairpin oligonucleotides. This provided their rank of order binding to all possible 5 base pair sequences and completely defined their DNA binding sequence selectivity. The technique, which conservatively requires manual measurement times of 1–2 min/plate (15 min/compound/512 hairpins) on a fluorescent plate reader reproduced the known properties of distamycin A and revealed distinctions with **128** responsible for the differences in binding the two 14 base pair androgen response elements. Combining the assay of a library of compounds against a library of DNA hairpin oligonucleotides, with automation of the assay, provides qualitative and/or quantitative information on the binding of all library members against a library of available sequences. Studies on extensions of this work are in progress and will be disclosed in due time.

Experimental Section

Methyl 4-[[(4-*tert*-Butyloxycarbonyl)amino-1-methylpyrrol-2-yl]carbonyl]amino-1-methylpyrrole-2-carboxylate (2). Initial conditions: a solution of **1a** (250 mg, 1.05 mmol, 1 equiv) and **1b** (200 mg, 1.05 mmol, 1 equiv) in DMF (5 mL) was treated with EDCI (403 mg, 2.1 mmol, 2 equiv) and DMAP

(320 mg, 2.6 mmol, 2.5 equiv) and the resulting solution was stirred for 14 h at 25 °C. The reaction mixture was poured into EtOAc (50 mL) and washed with 10% aqueous HCl (3 x 50 mL) and saturated aqueous $NaHCO_3$ (3 x 50 mL). The organic phase was dried (Na_2SO_4), filtered and concentrated to provide **2** (350 mg, 89%) as a tan foam. For optimized large scale: A solution of **1a** (3.8 g, 15.8 mmol, 1 equiv) and **1b** (3.0 g, 15.8 mmol, 1 equiv) in DMF (40 mL) and CH_2Cl_2 (10 mL) was treated with EDCI (4.5 g, 23.5 mmol, 1.5 equiv) and DMAP (2.3 g, 18.9 mmol, 1.2 equiv) and the resulting solution was stirred for 18 h at 25 °C. The reaction mixture was poured into EtOAc (60 mL) and washed with 10% aqueous HCl (3 x 50 mL) and saturated aqueous $NaHCO_3$ (3 x 50 mL). The organic phase was dried (Na_2SO_4), filtered and concentrated to provide **2** (5.8 g, 97%): mp 78–79 °C.

Methyl 4-[[[4-[[(4-*tert*-Butyloxycarbonyl)amino-1-methylpyrrol-2-yl]carbonyl]amino-1-methylpyrrol-2-yl]carbonyl]amino-1-methylpyrrole-2-carboxylate (3).

Initial conditions: a sample of **2** (50 mg, 0.13 mmol, 1 equiv) was treated with 4.0 N HCl/EtOAc (1 mL). The reaction mixture was stirred at 25 °C for 30 min, then concentrated and dried under reduced pressure for 1 h. EDCI (50 mg, 0.27 mmol, 2 equiv), DMAP (33 mg, 0.27 mmol, 2 equiv), and **1a** (63 mg, 0.27 mmol, 2 equiv) were added to a solution of the crude amine in DMF (1 mL). The reaction mixture was stirred for 3 h at 25 °C, diluted with EtOAc (10 mL) and washed with 10% aqueous HCl (3 x 10 mL) and saturated aqueous $NaHCO_3$ (3 x 10 mL). The organic phase was dried (Na_2SO_4), filtered and concentrated to a yellow solid. The solid material was suspended in 1:1 MeOH/10% NaOH (20 mL) and stirred for 30 min at 25 °C to decompose small amounts of contaminate symmetrical anhydride. The solution was then poured into EtOAc (20 mL) and washed with $NaHCO_3$ (3 x 20 mL). The organic phase was dried (Na_2SO_4), filtered and concentrated to provide **3** (47 mg, 73%) as a yellow foam: For optimized large scale: dipyrrole **2** (2.8 g, 7.43 mmol, 1 equiv) was treated with 4.0 N HCl/EtOAc (20 mL). The reaction mixture was stirred at 25 °C for 30 min, then concentrated and dried under reduced pressure. EDCI (2.1 g, 11.1 mmol, 1.5

equiv), DMAP (1.1 g, 9.0 mmol, 1.2 equiv), and **1a** (2.0 g, 8.2 mmol, 1.1 equiv) were added to a solution of the crude amine in DMF (100 mL). The reaction mixture was stirred for 3 h at 25 °C, diluted with EtOAc (100 mL) and washed with 10% aqueous HCl (3 x 100 mL) and saturated aqueous $NaHCO_3$ (3 x 100 mL). The organic phase was dried (Na_2SO_4), filtered and concentrated to provide **3** (3.6 g, 96%) as a yellow foam: mp 131–133 °C.

4-[[[4-[[(4-*tert*-Butyloxycarbonyl)amino-1-methylpyrrol-2-yl]carbonyl]amino-1-methylpyrrol-2-yl]carbonyl]amino-1-methyl-2-(((carbonyl)amino)propio-3-nitrile)pyrrole (4).

Tripyrrole **3** (70 mg, 0.14 mmol, 1 equiv) in THF/MeOH (3:1, 2 mL) was treated with a solution of LiOH (24 mg, 0.56 mmol, 4 equiv) in H_2O (0.5 mL). The solution was warmed at 60 °C for 5 h then diluted with EtOAc (20 mL) and H_2O (20 mL). The layers were separated and the aqueous layer was brought to pH 3 with 10% aqueous HCl. The resulting slurry was extracted with EtOAc (4 x 20 mL) and the combined organic extracts were dried (Na_2SO_4), filtered and concentrated. The crude acid (30 mg, 0.062 mmol) in DMF (1 mL) was treated with EDCI (23 mg, 0.12 mmol, 2 equiv) and DMAP (19 mg, 0.16 mmol, 2.5 equiv), followed by 3-aminopropionitrile (13 mg, 0.12 mmol, 2 equiv). The reaction mixture was stirred for 14 h at 25 °C, then diluted with EtOAc (20 mL) and washed with 10% aqueous HCl (3 x 20 ml) and saturated aqueous $NaHCO_3$ (3 x 20 mL). The organic phase was dried (Na_2SO_4), filtered and concentrated to afford 4 (31 mg, 95%) as a yellow solid: mp 170–172 °C.

Distamycin A Hydrochloride.

A solution of nitrile **4** (12 mg, 0.022 mmol) in dry EtOH (0.3 mL) was treated with 8.0 N HCl/EtOH (1 mL) at 0 °C for 30 min, then slowly warmed to 25 °C and stirred for 2 h. The solvent was removed under a stream of N_2 and the residue was washed with Et_2O (3 mL) and dried in vacuo for 30 min. The resulting solid was taken up in EtOH (0.3 mL) and treated with 7% NH_3/EtOH (1 mL) at 25 °C. After 1 h the reaction was concentrated to a tan solid and dried under reduced pressure for 1 h. The crude amidine was dissolved in MeOH (0.2 mL) and cooled to –40 °C. The solution was treated with a solution containing *N*-

formylimidazole, prepared by treating carbonyldiimidazole (18 mg, 0.11 mmol) in THF (0.4 mL) with a solution of formic acid (4.3 mL, 0.11 mmol) in THF (0.4 mL) at 25 °C for 15 min. The reaction mixture was stirred at –40 °C for 1 h, then concentrated to a volume of 0.2 mL. The product was precipitated with EtOAc (1 mL), and collected by filtration. The crude product was dissolved in cold *i*-PrOH (2 mL) containing decolorizing carbon (100 mg), stirred at 0 °C for 30 min, filtered and concentrated to a light yellow solid. The solid material was taken up EtOAc/acetone/MeOH/0.01 N HCl (5:3:1:1, 2 mL) and stirred with SiO_2 for 30 min, then filtered through Celite to remove traces of NH_4Cl and afford pure distamycin A (4.9 mg, 45%) identical in all respects with authentic material: mp 186–188 °C.

Ethidium Bromide Assay.

DNA hairpin oligonucleotides were purchased from Genbase Inc. (San Diego) as 880 μM (base pairs) solutions in water and stored as stock solutions at –80 °C. Prior to use, each oligonucleotide was diluted to 88 μM in water and stored at 0 °C for no longer than two days. Each well of a 96-well plate was loaded with Tris buffer containing ethidium bromide (0.1 M Tris, 0.1 M NaCl, pH 8, 0.44×10^{-5} M ethidium bromide final concentration, 88 μL). To each well was added one hairpin oligonucleotide (10 μL, 0.88×10^{-5} M in DNA base pairs final concentration). To each well was added distamycin A (2 μL of a 0.1 mM solution in water, 2.0×10^{-6} M final concentration) or **128** (6 μL of a 0.1 mM solution in water, 6.0×10^{-6} M final concentration). After incubation at 25 °C for 30 min, each well was read on a fluorescent plate reader (ex. 545 nm, em. 595 nm, cutoff filter at 590 nm) in duplicate experiments with two control wells (no distamycin = 100% fluorescence, no DNA = 0% fluorescence). Fluorescence readings are reported as % fluorescence relative to the controls. In our experience, fluorescence plate readers show a variability of ±10%, but surface effects (i.e bubbles, dust) may contribute to larger variations requiring a second set of measurements.

Determination of Distamycin A Binding Constants with Hairpin Oligonucleotides.

A 3 mL quartz cuvette was loaded with Tris buffer (0.1 M Tris, 0.1 M NaCl,

pH 8) and ethidium bromide (0.44 x 10^{-5} M final concentration). The fluorescence was measured (ex. 545 nm, em. 595 nm) and normalized to 0% relative fluorescence (free ethidium bromide is only weakly fluorescent). The DNA hairpin oligonucleotide of interest was added (0.88 x 10^{-5} M in DNA base pairs final concentration), the fluorescence was measured again and normalized to 100% relative fluorescence. A solution of distamycin A (1 µL, 0.1 mM in DMSO) was added and the fluorescence was measured following 30 min of incubation at 23 °C (each measurement was conducted four times and averaged). The addition of 1 µL aliquots was continued until the relative fluorescence had decreased to ≤ 50%. Binding constants (*K*) were calculated from $K = 1/[\text{agent}]-0.5[\text{DNA}]r$ where *K* = binding constant, [agent] = concentration at 50% reduction in fluorescence, [DNA] = DNA concentration, and r = 1/binding site size, using r = 1/8 or 0.125.

Preparation of Positional Scanning Libraries:

Herein we report the preparation of two comparable positional scanning libraries (Houghten, R. A.; et al. *Nature* **1991**, *354*, 84; Houghten, R. A.; et al. *BioTechniques* **1992**, *13*, 412; Pinilla, C.; et al. *BioTechniques* **1992**, *13*, 901; Dooley, C. T.; Houghten, R. A. *Life Sci.* **1993**, *52*, 1509; Smith, P. W.; et al. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2821; Pirrung, M. C.; Chen, J. *J. Am. Chem. Soc.* **1995**, *117*, 1240) composed of 1000 members each and the results of their comparison examination.

Unlike solid-phase synthesis where polymer bound substrate is the stoichiometry limiting reaction partner, either the substrate or the reacting attachment group may be limiting in solution-phase chemistry. This dictates the use of mix and split synthesis for the solid-phase in order to accommodate differential reaction rates, whereas the simpler protocol of mixture synthesis with limiting reagent stoichiometry may be used in solution to ensure all library members are generated (The exception for solid-phase synthesis enlists an excess of the reacting monomers in adjusted concentrations to accommodate the different reaction rates and requires that this relative rate information be available at the onset of the mixture synthesis. Houghten, R. A.; et al. *Nature* **1991**, *354*, 84; Houghten, R. A.; et al. *BioTechniques* **1992**, *13*, 412; Pinilla, C.; et al.

BioTechniques **1992**, *13*, 901; Dooley, C. T.; Houghten, R. A. *Life Sci.* **1993**, *52*, 1509). The implementation of the latter only requires the ability to remove unreacted starting substrate. Although this is not possible with solid-phase synthesis, this can be accomplished by aqueous acid/base extractions (Cheng, S.; et al. *J. Am. Chem. Soc.* **1996**, *118*, 2567; Boger, D. L.; et al. *J. Am. Chem. Soc.* **1996**, *118*, 2109; Cheng, S.; et al. *Bioorg. Med. Chem.* **1996**, *4*, 727; Boger, D. L.; et al. *Bioog. Med. Chem. Lett.* **1997**, *7*, 1903; Boger, D. L.; et al. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 463; Boger, D. L.; Chai, W. *Tetrahedron* **1998**, *54*, 3955; Boger, D. L.; et al. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2339; Boger, D. L.; et al. *Bioorg. Med. Chem.* **1998**, *6*, 1347; Boger, D. L.; et al. *J. Org. Chem.* **1999**, *64*, 7094; Boger, D. L.; et al. *Helv. Chim. Acta* submitted) in the work we describe, which also serves to remove reactants, reagents, and reagent byproducts providing clean products.

The synthesis of positional scanning libraries represents one of the most useful protocols for mixture synthesis. Not only is it much less time intensive than the parallel synthesis of individual compounds or small mixtures, but it produces depository libraries for use in multiple screens with immediate deconvolution (Houghten, R. A.; et al. *Nature* **1991**, *354*, 84; Houghten, R. A.; et al. *BioTechniques* **1992**, *13*, 412; Pinilla, C.; et al. *BioTechniques* **1992**, *13*, 901; Dooley, C. T.; Houghten, R. A. *Life Sci.* **1993**, *52*, 1509; Geysen, H. M.; et al. *Mol. Immunol.* **1986**, *23*, 709; Erb, E.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11422; Deprez, B.; et al. *J. Am. Chem. Soc.* **1995**, *117*, 5405; Boger, D. L.; et al. *J. Am. Chem. Soc.* **1998**, *120*, 7220; Boger, D. L.; et al. *J. Org. Chem.* **2000**, *65*, 1467). Thus, unlike other deconvolution protocols (Geysen, H. M.; et al. *Mol. Immunol.* **1986**, *23*, 709; Erb, E.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11422; Deprez, B.; et al. *J. Am. Chem. Soc.* **1995**, *117*, 5405; Boger, D. L.; et al. *J. Am. Chem. Soc.* **1998**, *120*, 7220; Boger, D. L.; et al. *J. Org. Chem.* **2000**, *65*, 1467), positional scanning libraries provide lead identities in a single round of testing. Despite these attributes, it was not clear how well such libraries would perform in screening for DNA binding agents relative to other formats (Freier, S. M.; et al. *J. Med. Chem.* **1995**, *38*, 344; Konings, D. A. M.; et al. *J. Med. Chem.* **1997**, *40*, 4386).

Library Design and Synthesis:

In order to insure that the quality of information derived from the library assessment could be established, two 1000-member libraries were prepared that contain the same compounds assembled in our prior study (Boger, D. L.; et al. *J. Am. Chem. Soc.* **2000**, *122*, 0000). Each positional scanning library consists of 30 sublibraries that can be divided into three sets. These sets differ in the fixed positions of a monomer subunit within the tripeptide (Figure 1). Thus, the library was prepared by substitution of the same ten subunits for each of the three 4-aminopyrrole-2-carboxylic acid subunits of distamycin A (Figure 2). Included in this set was the authentic 4-aminopyrrole-2-carboxylic acid subunit of distamycin A, insuring that the natural product analogue was also among the library members. The C-terminus of the library compounds was capped as methyl or ethyl esters and the *N*-terminus was acylated with 4-dimethylaminobutyric acid (DMABA), a basic side chain that mimics the distamycin A amidine, providing analogues that bear functionalization and a substitution pattern established to provide DNA affinities comparable to that of the natural product (Boger, D. L.; et al. *J. Am. Chem. Soc.* **2000**, *122*, 0000).

The synthesis of the library was divided into four parts (Figure 3). First, a mixture of 100 dimers was synthesized on a 144 μmol scale by coupling the set of ten amino acid esters **1a** and **5a–13a** with the corresponding set of ten BOC-amino acids **1b** and **5b–13b** using 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride (EDCI) and dimethylaminopyridine (DMAP) as an additive. For the preparation of sublibraries I, where the first position within the trimer is fixed with a single **A** residue, ten portions of the dimer mixture were deprotected with HCl/EtOAc and coupled to ten individual BOC-amino acids providing ten sublibraries each containing a different and single **A** residue. The set of ten sublibraries II was assembled by coupling ten individual BOC-amino acids to a mixture of amino acid esters. Subsequent deprotection of the BOC-group and coupling to a mixture of BOC-amino acids yielded the set of ten trimer sublibraries each containing a single and different **B** residue. Finally, the dimer mixture of 100 compounds was saponified with LiOH, divided into ten portions, and coupled with ten individual amino acid esters (**C** residue) to give the third set of sublibraries III. The entire library containing 1000 compounds was

synthesized conducting 43 reactions.

The 30 positional scanning sublibraries were also converted into their corresponding dimethylaminobutyric acid (DMABA) derivatives as shown in Figure 4. Yields of the BOC- and DMABA-trimer libraries are given in Figure 5.

Cytotoxic Activity:

Evaluation of the original library in a cellular functional assay for L1210 cytotoxic activity revealed two structurally-related BOC-trimers, BOC-A_9-B_9-C_5-OEt (**11**, 29 nM) and BOC-A_8-B_9-C_5-OEt (**12**, 68 nM), which exhibited uniquely potent activity (Boger, D. L.; et al. *J. Am. Chem. Soc.* **2000**, *122*, 0000). Both were identified in a single deconvolution of a potent sublibrary of ten compounds, which exhibited activity 10-fold more potent than any other sublibrary and >100 times more potent than 90% of the mixtures. While it cannot be excluded that additional unidentified members of the library exhibit comparable activity, the uniquely potent activity of the sublibrary containing **11** and **12** and the dependable performance of the small mixture testing reflecting the composite activity of the components suggest that **11** and **12** are at least 10 times more potent than any other library member. The testing of the 30 positional scanning libraries (Figures 6 and 7) also revealed the identity of **11** and **12** (Figure 8), but required the preparation of more candidate structures in the deconvolution of the activity, 16 compounds in total. Notably, both **11** and **12** are 1000 times more potent than distamycin A.

The most potent residues identified in the scanning library were A_9, B_9 and B_1, and C_{10}, C_2, C_5 and C_6. Importantly, the combination of the most potent residues, BOC-A_9-B_9-C_{10}-OEt, was not a compound that exhibited potent cytotoxic activity. Moreover, none of the alternative 14 possible combinations exhibited cytotoxic activity that approached that of **11** and **12**. Only **13** (BOC-A_9-B_1-C_5-OEt) exhibited respectable cytotoxic activity (IC_{50} = 0.42 μM) and this compound was still 15-fold less active than **11**.

The distinctions between the activities of the positional scanning libraries are small and smaller than those of the ten compound mixtures prepared by parallel synthesis (Boger, D. L.; et al. *J. Am. Chem. Soc.* **2000**, *122*, 0000). This is a consequence of testing mixtures of 100 compounds (positional scanning)

versus ten compounds (parallel synthesis) where the impact of any single compound in the mixture is diminished. Nonetheless, the two potent compounds in the library of 1000 were detected. Thus, the ease of the synthesis of the initial positional scanning library relative to the small mixture library assembled by the more time consuming parallel synthesis is offset by the less distinct biological discrimination observed in the assay of the library and the increased effort required in the deconvolution.

The evaluation of the second positional scanning DMABA-trimer library provided observations that were analogous to those detected with the small mixtures prepared by parallel synthesis. In general, the DMABA-trimers were less active than the corresponding BOC-trimers (Figure 9). The exceptions, as detected in the previous studies with the small mixtures, tend to be the residues which convey insolubility to the resulting compounds (e.g. C_7). Presumably, the protonated side chain of the DMABA-trimers offsets this insolubility contributing productively to characteristics that enhance their activity. Only one DMABA-trimer mixture was deconvoluted in our prior work (Boger, D. L.; et al. *J. Am. Chem. Soc.* **2000**, *122*, 0000). Although it was the most potent, ten mixtures exhibited comparable activity, $IC_{50} < 1$ but > 0.1 μM. Aside from deconvolution of the most potent of these mixtures, $IC_{50} = 0.42$ μM, no effort was made to deconvolute the remaining mixtures, even though they all would contain compounds with similar activities. This deconvolution provided DMABA-A_8-B_4-C_7-OMe (**14**) with an $IC_{50} =$ 0. 46 μM. The examination of the second positional scanning library did not suggest this compound as a candidate lead (Figure 9). The most potent residues identified were A_9 and A_5, B_8, and C_{10}. The preparation and testing of the two candidate structures DMABA-A_9-B_8-C_{10}-OEt (**15**) and DMABA-A_5-B_8-C_{10}-OEt (**16**) revealed IC_{50}`s of 0.32 μM and 3.2 μM, respectively. Thus, although **14** was not identified, a structure (**15**) of comparable activity was identified. Notably, both **14** and **15** are 100 times more potent than distamycin A ($IC_{50} = 42$ μM), and its close tripyrrole analogue **49** ($IC_{50} = 44$ μM) (Figure 10).

DNA Binding Properties:

The positional scanning DMABA-trimer libraries were screened enlisting the ethidium bromide displacement DNA binding assay (Boger, D. L.; et al. *J.*

Am. Chem. Soc. **2000**, *122*, 0000; Morgan, A. R.; et al. *Nucleic Acids Res.* **1979**, *7*, 547; Baguley, B. C.; Falkenhaug, E.-M. *Nucleic Acids Res.* **1978**, *5*, 161; Boger, D. L.; et al. *Chem.-Biol. Interact.* **1990**, *73*, 29; Boger, D. L. Sakya, S. M. *J. Org. Chem.* **1992**, *57*, 1277) with two hairpin oligonucleotides that were used in our prior study (Boger, D. L.; et al. *J. Am. Chem. Soc.* **2000**, *122*, 0000). They constitute the dimer androgen receptor binding consensus sequences, ARE-consensus (Cato, A. C. B.; et al. *EMBO J.* **1987**, *33*, 545) and PSA-ARE-3 (Cleutjens, K. B. J. M.; et al. *Mol. Endocrinol.* **1993**, *7*, 23), and targets for chemotherapeutic resistant prostate cancer (Chang, C. C.; et al. *Crit. Rev. Eucaryotic Gene Expression* **1995**, *5*, 97; Bentel, J. M.; Tilley, W. D. *J. Endocrinology* 1996, *151*, 1; Veldscholte, J.; et al. *Biochem. Biophys. Res. Commun.* **1990**, *173*, 534; Galbraigth, S. M.; Duchesne, G. M. *Eur. J. Cancer* **1997**, *33*, 545). The latter contains a 5 base-pair AT-rich sequence known to bind distamycin A, while the former contains the same sequence interrupted by a single GC base-pair. The screening of the library, which entails measurement of the loss of fluorescence derived from compound binding and displacement of prebound ethidium bromide, identified A_6, B_6, and C_6 as the most effective residues for binding to the PSA-ARE-3 hairpin containing the 5 base-pair AT-rich site as well as the ARE-consensus hairpin. However, binding to the latter sequence was less effective (Figure 11). This constitutes the identification of DMABA-A_6-B_6-C_6-OMe (**49**), the direct distamycin A analogue, in the 1000-member library as the most effective agent. This successful identification of **17** from the positional scanning library is tempered by the fact that it did not identify as candidate binders DMABA-A_8-B_9-C_5-OEt (**230**) or DMABA-A_{10}-B_8-C_6-OMe (**240**), which were identified in our prior study (Boger, D. L.; et al. *J. Am. Chem. Soc.* **2000**, *122*, 0000). In addition, since the decreases in affinity for binding to the GC interrupted ARE-consensus hairpin were rather small and uniform, the ability to detect compounds including **240** that bound both sequences (Boger, D. L.; et al. *J. Am. Chem. Soc.* **2000**, *122*, 0000) equally well was not possible.

This is a natural consequence of the testing of the larger 100 compound mixtures and the relative insensitivity of the assay to the contribution of any single, uniquely acting compound in the mixture. Thus, the more global

observations are accurately detected with the positional scanning library and a useful lead structure with defined properties was identified. However, more subtle discoveries within the library were not identified. Thus, the disadvantages associated with the loss of their detection and this information contained within the library must be balanced against the advantages of the ease of synthesis of the parent libraries and judged in light of the objectives of the library synthesis. Typically, the positional scanning libraries would be most effective for lead identification and would be less suitable for lead optimization.

Experimental

Synthesis of dimer mixtures.

Solutions of a mixture of the ten BOC-protected acids **1b** and **5–13b** (each 160 μmol, 1.1 equiv) and the ten amino esters **1a** and **5–13a** (each 144 μmol, 1 equiv) in DMF (20 mL) were treated with EDCI (3.2 mmol, 22.2 equiv) and DMAP (3.6 mmol, 25.0 equiv). The resulting solutions were stirred for 12–16 h, then the DMF was removed under reduced pressure and the resulting oil was taken up in EtOAc (20 mL) and washed with 10% aqueous HCl (3 × 20 mL) and saturated aqueous $NaHCO_3$ (3 × 20 mL). The resulting solutions were dried (Na_2SO_4), filtered, and concentrated to dryness providing the mixture of dimers that was used without purification.

General procedure for preparation of sublibraries I.

BOCNH-A_{1-10}-CONH-X-CONH-X-OR. Ten individual portions of the mixture of dimers BOCNH-X-CONH-X-OR (144 μmol, 1.0 equiv) were dissolved in 4.0 N HCl/EtOAc (1 mL), and the mixtures were stirred at 25 °C for 2 h. The solvent was removed under a stream of N_2 and the residues were dried *in vacuo* for 4 h. Each sample was dissolved in DMF (1.0 mL) and treated with one of the ten BOC-carboxylic acids **1b** and **5–13b**, followed by EDCI (67.5 mg, 352 μmol, 2.2 equiv) and DMAP (48.9 mg, 400 μmol, 2.5 equiv). The solutions were stirred for 16 h at 25 °C. The mixtures were poured into EtOAc (20 mL) and washed with 10% aqueous HCl (3 × 20 mL), followed by saturated aqueous $NaHCO_3$ (3 × 20 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated *in vacuo*

to afford the final trimers (68–99%).

General procedure for preparation of sublibraries II.

BOCNH-X-CONH-B_{1-10}-CONH-X-OR. Ten single BOC-amino acids **1b** and **5–13b** (160 µmol, 11 equiv) and a mixture of ten amino acid esters **1a** and **5–13a** (each 14.4 µmol, 1 equiv) were dissolved in DMF (1.5 mL) and treated with EDCI (75 mg, 390 µmol, 24.4 equiv) and DMAP (49 mg, 400 µmol, 25 equiv). The solutions were stirred for 16 h at 25 °C. The mixtures were poured into EtOAc (20 mL) and washed with 10% aqueous HCl (3 × 20 mL), followed by saturated aqueous $NaHCO_3$ (3 × 20 mL). The organic phases were dried (Na_2SO_4), filtered, and concentrated in vacuo to afford ten mixtures of dipeptides BOCNH-B_{1-10}-CONH-X-OR. Each of these ten mixtures was dissolved in 4.0 N HCl/EtOAc (1 mL), and the mixtures were stirred at 25 °C for 2 h. The solvent was removed under a stream of N_2 and the residues were dried *in vacuo* for 4 h. Each sample was dissolved in DMF (1.5 mL) and treated with a mixture of the ten BOC-carboxylic acids **1b** and **5–13b** (each 15.0 µmol, 1.04 equiv) followed by EDCI (69 mg, 360 µmol, 24 equiv) and DMAP (44 mg, 360 µmol, 24 equiv). The solutions were stirred for 16 h at 25 °C. The mixtures were poured into EtOAc (20 mL) and washed with 10% aqueous HCl (3 × 20 mL), followed by saturated aqueous $NaHCO_3$ (3 × 20 mL). The organic phases were dried (Na_2SO_4), filtered, and concentrated in vacuo to afford the final sublibraries II (68–99%).

General procedure for preparation of sublibraries III.

BOCNH-X-CONH-X-CONH-C_{1-10}-OR. The mixture of dimers BOCNH-X-CONH-X-OR (190 µmol, 1.0 equiv) were dissolved in THF/MeOH/H_2O (20 mL, 2:1:1), LiOH (760 µmol, 4 equiv) was added and the mixture was stirred at 25 °C for 18 h. The solvent was removed under reduced pressure and the residue was acidified with 10% aqueous HCl. The dipeptides acids were extracted with EtOAc (3 × 20 mL), the combined organic layers were washed with 10% aqueous HCl and water, dried (Na_2SO_4), filtered, and concentrated *in vacuo*. Ten individual portions of the mixture of BOCNH-X-CONH-X-OH (182 µmol, 1.14 equiv) were dissolved in DMF (2.0 mL) and treated with one of the ten amino acid esters **1a** and **5–13a** (each 160 µmol, 1 equiv) followed by EDCI (76.7 mg, 400

μmol, 2.5 equiv) and DMAP (48.9 mg, 400 μmol, 2.5 equiv). The solutions were stirred for 16 h at 25 °C. The mixtures were poured into EtOAc (20 mL) and washed with 10% aqueous HCl (3 × 20 mL), followed by saturated aqueous $NaHCO_3$ (3 × 20 mL). The organic phases were dried (Na_2SO_4), filtered and concentrated *in vacuo* to afford the final sublibraries III (45–81%).

General procedure for preparation of DMABA-trimer libraries. Each of the BOC-trimer sublibraries (0.007 mmol, 1 equiv) was dissolved in 4.0 N HCl/EtOAc (1 mL), and the mixtures were stirred at 25 °C for 2 h. The solvent was removed under a stream of N_2 and the residues were dried *in vacuo* for 4–8 h. Each sample was treated with dimethylaminobutyric acid as a 0.1 M solution in DMF (150 μL, 0.015 mmol, 2 equiv) followed by EDCI as a 0.1 M solution in DMF (180 μL, 0.018 mmol, 2.5 equiv) and DMAP as a 0.1 M solution in DMF (219 μL, 0.022 mmol, 3.0 equiv). The solutions were stirred for 12–16 h at 25 °C. The solvent was removed under a stream of N_2 and the residues were taken up in H_2O (10 mL) and extracted with EtOAc (4 × 10 mL). The combined organic layers were dried (Na_2SO_4), filtered, and concentrated *in vacuo*. The resulting solids were slurried in Et_2O (1 mL) and centrifuged. The pellets were again slurried in Et_2O (1 mL) and collected by filtration to afford the desired DMABA-trimers (22–100%).

General procedure for preparation of individual BOC-trimers.

The individual dipeptides BOCNH-X-CONH-Y-OR (X = **5**, **12**; Y = **6**, **9**, **1**, **13**) (1.0 equiv) were dissolved in 4.0 N HCl/EtOAc (1 mL), and the mixtures were stirred at 25 °C for 2 h. The solvent was removed under a stream of N_2 and the residues were dried *in vacuo* for 4 h. Each sample was dissolved in DMF (1.0 mL) and was treated with a BOC-carboxylic acid (**11b**, **12b**), followed by EDCI (2 equiv) and DMAP (2.5 equiv). The solutions were stirred for 16 h at 25 °C. The mixture was poured into EtOAc (10 mL) and washed with 10% aqueous HCl (3 × 10 mL), followed by saturated aqueous $NaHCO_3$ (3 × 10 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated *in vacuo* to afford the final trimers (37–99% yield).

BOCNH-9-CONH-1-CONH-10-OEt. (9.1 mg, 79%);

BOCNH-9-CONH-1-CONH-2-OMe. (9.2 mg, 99%);
BOCNH-9-CONH-1-CONH-6-OMe. (9.0 mg, 70%);
BOCNH-9-CONH-1-CONH-5-OEt (13). (8.7 mg, 70%);
BOCNH-8-CONH-9-CONH-10-OEt. (19.3 mg, 81%);
BOCNH-8-CONH-9-CONH-2-OMe. (20.3 mg, 77%);
BOCNH-8-CONH-9-CONH-6-OMe. (18.2 mg, 71%);
BOCNH-8-CONH-1-CONH-10-OEt. (8.2 mg, 49%);
BOCNH-8-CONH-1-CONH-2-OMe. (17.4 mg, 99%);
BOCNH-8-CONH-1-CONH-6-OMe. (10.9 mg, 82%);
BOCNH-8-CONH-1-CONH-5-OEt. (4.8 mg, 37%).

General procedure for preparation of individual DMABA-trimers.

Each of the individual samples of BOCNH-X-CONH-8-CONH-10-OEt trimers (X = **9**, **12**) (1.0 equiv) was dissolved in 4.0 N HCl/EtOAc (1 mL), and the mixtures were stirred at 25 °C for 2 h. The solvent was removed under a stream of N_2 and the residues were dried in vacuo for 4 h. Each sample was treated with 4-dimethylaminobutyric acid (2 equiv), EDCI (2 equiv), DMAP (2.5 equiv), and DMF (1 mL). The solutions were stirred for 16 h at 25 °C. The solvent was removed under a stream of N_2 and the residues were taken up in H_2O (10 mL) and extracted with EtOAc (4 × 10 mL). The combined organic layers were dried (Na_2SO_4), filtered, and concentrated *in vacuo*. The resulting solids were slurried in Et_2O (10 mL) and centrifuged. The pellets were again slurried in Et_2O (10 mL) and collected by filtration to afford the desired DMABA-trimers (51–80% yield).
$Me_2NCH_2CH_2CH_2CONH$-9-CONH-8-CONH-10-OEt (15). (18.9 mg, 80%);
$Me_2NCH_2CH_2CH_2CONH$-5-CONH-8-CONH-10-OEt (16). (2.2 mg, 51%).

Ethidium bromide displacement assays.

Hairpin oligonucleotides (0.887×10^{-5} M bp) were mixed with ethidium bromide (0.444×10^{-5} M) in a 2:1 ratio of base-pair:ethidium bromide in a 0.1 M Tris–HCl, 0.1 M NaCl, pH 8 buffer. The fluorescence measurements were conducted at 545 nm excitation and 595 nm emission. For the rapid screening of libraries, 96-well plates (Costar: black, 360 µL, flat-bottom) were loaded with the premixed ethidium bromide/DNA solution (100 µL) and single aliquots of each

library (1 μL of 10 mM solutions in DMSO, 99 μM final concentration) were added. Each plate was incubated at 25 °C for 30 min before reading on a fluorescence plate reader (Molecular Devices SpectraMax Gemini) using 545 nm excitation and 595 nm emission.

Detailed Description of Figures:

Figure 1 shows how the positional scanning libraries were designed. Each positional scanning library consists of 30 sublibraries that can be divided into three sets. These sets differ in the fixed positions of a monomer subunit within the tripeptide. Some of the same compounds which were assembled in a prior study were in the two 1000-member libraries. This was to insure that the quality of information derived from the library assessment could be confirmed.

Figure 2 shows the structures of amino acid monomer units used in the preparation of the libraries. The library was prepared by substitution of the same 10 subunits for each of the three 4-aminopyrrole-2-carboxylic acid subunits of distamycin A. Included in this set was the authentic 4-aminopyrrole-2-carboxylic acid subunit of distamycin A so that the natural product analogue was also among the library members. The C-terminus of the library compounds was capped as methyl or ethyl esters and the N-terminus was acylated with 4-dimethylaminobutyric acid (DMABA), a basic side chain that mimics the distamycin A amidine, providing analogues which bear functionalization and a substitution pattern established to provide DNA affinities comparable to that of the natural product. Another set of library compounds had the N-terminus capped with a *tert*-butyloxycarbonyl group which renders it neutral and non-nucleophilic.

Figure 3 is a scheme illustrating how the positional scanning libraries were synthesized. The synthesis of the library was divided into four parts. First, a mixture of 100 dimers was synthesized on a 144 mmol scale by coupling the set of ten amino acid esters **1a, 5a-13a** with the corresponding set of ten BOC-amino acids **1b, 5b-13b** using 1-[3-(dimethylamino)-propyl]-3-ethylcarbodiimide hydrochloride (EDCI) and dimethylaminopyridine (DMAP) as an additive. For the preparation of sublibraries I, where the first position within the trimer is fixed with a single **A** residue, ten portions of the dimer mixture were deprotected with HCl/EtOAc and coupled to ten individual BOC-amino acids providing 10

sublibraries each containing a different and single **A** residue. The set of ten sublibraries II was assembled by coupling ten individual BOC-amino acids to a mixture of amino acid esters. Subsequent deprotection of the BOC-group and coupling to a mixture of BOC-amino acids yielded the set of 10 trimer sublibraries each containing a single and different **B** residue. Finally, the dimer mixture of 100 compounds was saponified with LiOH, divided in ten portions, and coupled with ten individual amino acid esters (**C** residue) to give the third set of sublibraries III.

Figure 4 is a scheme showing how the three sets of ten libraries were functionalized on the N-terminus. The 30 positional scanning libraries were also converted into their corresponding dimethylaminobutyric acid (DMABA) derivatives as shown in Scheme 2.

Figure 5 is a table showing the yields of the BOC- and DMABA-trimers.

Figure 6 is a bar graph showing the most potent residues that were found using the positional scanning library. The most potent residues identified in the scanning library were A_{11} and A_{12}, B_{12} and B_5, and C_{13}, C_6, C_9 and C_1. The combination of the most potent residues, BOC-A_{12}-B_{12}-C_{13}-OEt, was not a compound that exhibited potent cytotoxic activity. Moreover, none of the remaining 14 possible combinations exhibited cytotoxic activity that approached that of **67** and **66**. Only **200** (BOC-A_{12}-B_5-C_9-OEt) exhibited respectable cytotoxic activity (IC_{50} = 0.42 mM) and this compound was still 15-fold less active than **67**.

Figure 7 is a table showing the cytotoxic activities of the candidate compounds composed of the most potent residues found by the positional scanning library. Only two compounds, **67** and **66**, exhibited potent cytotoxic activity while the third most potent compound was 15-fold less active than **67**.

Figure 8 shows the structures of the two most potent compounds found in the BOC-trimer library. Both of these compounds have the B_{12} and the C_9 residues in common.

Figure 9 shows the bar graph of the results of the positional scanning DMABA-trimer library. In general, the DMABA-trimers were less active than the corresponding BOC-trimers. The most potent residues identified were A_{12} and A_9, B_{11} and C_{13}. The preparation and testing of two candidate structures DMABA-A_{12}-B_{11}-C_{13}-OEt (**210**) and DMABA-A_9-B_{11}-C_{13}-OEt (**220**) revealed IC_{50}'s of 0.32 mM and 3.2 mM, respectively.

Figure 10 shows the structures of **86, 210, 220** and **49**. Compounds **86, 210** and **220** were identified as potent DMABA-trimers. Compound **49** is a close analogue of distamycin A. Compound **49** and distamycin A have IC_{50}'s of 42 mM and 44 mM respectively. Compound **86** was identified in a previous deconvolution study and has an IC_{50} = 0.46 mM which is in the range of activity of **210** and **220**.

Figure 11 shows the results of the positional scanning library for the ethidium bromide displacement assay with two hairpin nucleotides. The two hairpin nucleotides are part of the dimer androgen receptor binding consensus sequences, ARE-consensus and PSA-ARE-3. The screening of the library, which entails measurement of the loss of fluorescence derived from compound binding and displacement of prebound ethidium bromide, identified A_1, B_1 and C_1 as the most effective residues for binding to the PSA-ARE-3 hairpin containing the 5 base-pair AT-rich site as well as the ARE-consensus hairpin. Binding to the latter sequence was less effective. This constitutes the identification of DMABA-A_1-B_1-C_1-OMe (**49**), the direct distamycin A analogue, in the 1000 member library as the most effective binding agent.

Figure 12 illustrates the solution-phase strategy for developing libraries of new DNA binding agents. The strategy involves systematically replacing the N-methylpyrrole subunit with other heterocyclic amino acids to give a first generation library in a small mixture format. This is done by using liquid-liquid purification protocols. Then a basic side chain is added to expand the possible number of compounds and to mimic the amidine side chain of distamycin A.

Figure 13 is a simplified illustration of the general procedure for the rapid DNA binding screen. The chosen DNA as homopolymers, heteropolymers or hairpin oligonucleotides is placed in 96 well plates. Upon treatment with ethidium bromide there is a large increase in fluorescence as ethidium bromide intercalates with the DNA. When a non-fluorescent DNA binding agent is added there is a percentage decrease in the fluorescence due to binding. The percentage decrease in fluorescence is proportional to the extent of DNA binding. This provides the relative DNA binding affinities and through quantitative titration that may be carried out later, an accurate, absolute binding constant is obtained.

Figure 14 is a scheme showing the synthesis of distamycin A. Starting with the pyrrole carboxylic acid **1a**, (Baird, E. E.; Dervan, P. B. *J. Am. Chem.*

Soc. **1996**, *118*, 6141.) coupling with aminopyrrole **1b** (Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141.) using EDCI/DMAP afforded **2** in high yield (97%). Removal of the BOC protecting group with HCl/EtOAc followed by coupling to pyrrole **1a** afforded the tripeptide **3** in good yield (96%). Saponification of **3** followed by coupling with b-aminopropionitrile afforded nitrile **4** in excellent yield (95%). Treatment of nitrile **4** with HCl/EtOH followed by NH_3/EtOH afforded the desired amidine with concomitant removal of the BOC group. Due to the intrinsic instability of this free amine, it was immediately treated with *N*-formyl imidazole to afford distamycin A. This provided distamycin A in 40% overall yield for eight steps without deliberate optimization, and required only acid/base liquid–liquid extraction to afford all intermediates and the final product with >95% purity as demonstrated by their ^{1}H NMR spectra.

Figure 15 shows how two prototypical libraries of potential DNA binding agents were prepared in a small mixture format. Using eleven *N*-BOC heterocyclic amino acids and twelve amino esters, the individual subunits were coupled using EDCI/DMAP to provide all possible 132 individual dipeptides in parallel. The use of EDCI and DMAP allows for the removal of excess coupling agents and their reaction by-products along with unreacted starting materials by acid/base liquid–liquid extraction. These individual dimers were deprotected and coupled to a mixture of ten *N*-BOC carboxylic acids to give 132 mixtures of ten *N*-BOC-trimers where only the last position (subunit A) is undefined (1320 compounds). Removal of the BOC group and coupling to the basic side chain, *N, N*-dimethylaminobutyric acid (DMABA), affords an analogous DMABA-trimer library (1320 compounds).

Figure 16 shows the heterocyclic amino acids selected for the first prototypical libraries. This set includes the pyrrole, imidazole, and thiazole amino acids which were studied by Dervan and Lown and the indole and CDPI amino acids previously studied by the inventor.

Figure 17 shows the preparation of the dimers using **1b** and **5b-13c** as the acid component and **1a** and **5a**–**15a** as the amine component, 120 individual dimers were prepared. Each dimer was prepared in 70–80 mg quantities in parallel, using only acid/base liquid–liquid extraction purification to afford products in typically >80% yield and >95% purity.

Figure 18 shows the three instances when using the indole subunit **13b** as the acid component in couplings with the three unreactive amines (**5a**, **7a**, and **14a**), the diketopiperazine **30** was isolated due to indole dimerization. To circumvent this problem, the indole nitrogen was protected with a *p*-methoxybenzyl group to afford indole **31**. Hydrolysis to afford the free acid **32** and coupling to the three individual amines afforded the desired dimers in moderate yield. Simultaneous deprotection of both the *p*-methoxybenzyl and BOC-protecting groups (TFA/anisole; 60 °C) afforded the desired amines. Dimers could not be prepared where benzimidazole was the acid component. This monomer was only used in the third position (C) of the trimers.

Figure 19 shows how the preparation of the trimer libraries was investigated initially by preparing several sets of individual trimers to ensure the reaction conditions were appropriate. Deprotection of the dimer with HCl/EtOAc followed by coupling with **1b**, **5b**–**13b** afforded the ten BOC-trimers in high yield and with >90% purity using only acid/base liquid–liquid extraction purification. This set based on the dipyrrole dimer is of special interest because it contains a close analog of distamycin, the tripyrrole **39** (BOCNH-**1**-CONH-**1**-CONH-**1**-OMe).

Figure 20 shows the reaction conditions for adding the dimethylaminobutyric acid side chain (DMABA) to the individual trimers. The yields were much more variable than the previous steps since some of the derivatives were appreciably soluble. The typical acid/base liquid-liquid purification protocol was modified because of this. The solvent was removed from the reactions, the products were suspended in water and extraction with EtOAc gave the desired products.

Figure 21 shows the graphical results of the cytotoxicity assay (L1210) for the BOC-trimer libraries. Thirteen of the libraries showed activity at less than 1 μM, one showed activity at 100 nM. The most active library contained the benzofuran subunit (**12**) at the central position (subunit B) and the imidazole subunit (**9**) at the final position (subunit C).

Figure 22 shows the deconvolution of the mixture by the resynthesis of the 10 components, beginning with the stored BOCNH-**12**-CONH-**9**-OEt dimer. A second round of testing revealed that the most active components contained either the benzofuran (**12**) or the benzothiophene (**11**) at the first (A) position, with

the IC_{50}'s of 29 nM and 68 nM for **66** and **67**, respectively. When compared with distamycin A (IC_{50} = 42 μM), both are 1000 times more potent.

Figure 23 shows the graphical results of the cytotoxicity assay (L1210) for the DMABA-trimer libraries. The IC_{50} values for the DMABA-trimer libraries are on the order of 10-100 fold higher than the BOC-trimer libraries (Figure 10). The most active mixture contains the CDPI subunit (**10**) in the final position (subunit C), and the thiophene subunit (**8**) at the central position (subunit B).

Figure 24 illustrates the synthesis of the individual members of the library for deconvolution of the mixture. A second round of screening revealed that the most active component of this library contained the benzothiophene (**11**) at the first position (subunit A), with an IC_{50} of 0.46 μM for **86** and it was >10 times more active than any other compound in the mixture and 100 times more potent than any of the individual components of the **49-58** mixture based on and including the close distamycin analogues (Figure 12). The corresponding BOC-trimers were tested as well and these show a 10-100 fold greater activity than the DMABA-trimers.

Figure 25 shows the general procedure for establishing DNA binding of a library of compounds with a single sequence.

Figure 26 shows the binding results for the DMABA-trimer library with poly[dA]–poly[dT]. There were several general trends: (1) all the DMABA-trimers induce some decrease in fluorescence, indicating the libraries have an overall AT affinity; (2) high affinity libraries contain one of the larger subunits at the second (B) position (monomers **10–14**); and (3) the smaller subunits at the third (C) position (monomers **1, 5–9**) appear to be more active. Notably, four of the 10 compound mixtures showed a higher affinity than the pyrrole sublibrary containing **49**, the tripyrrole analog of distamycin. The highest affinity mixture contains the CDPI subunit (**10**) at the second position and the imidazole subunit (**9**) at the third position. The second most effective mixture contains the benzothiophene (**11**) at the second position and the pyrrole (**1**) at the third position.

Figure 27 shows the synthesis of mixtures that were deconvoluted, both the BOC- and DMABA trimers.

Figure 28 tabulates the results of the L1210 assay on individual compounds synthesized for deconvolution. The DNA binding properties of three

sets of individual compounds are also shown in the table. The activity of the mixtures in the cell-based assay approximated that of the individual components and established the reliability of testing in the small mixture format for libraries.

Figure 29 is a bar graph illustrating the binding affinity to two related sequences of the androgen response element, the 14-base pair ARE-consensus (Cato, A. C. B.; Hernerson, D.; Ponta, H. *EMBO J.* **1987**, *33*, 545) and the PSA-ARE-3 (Cleutjens, D. B. J. M.; et al., *Mol. Endocrinol.* **1993**, *7*, 23) sequences in the ethidium bromide displacement assay.

Figure 30 displays the same results of the ethidium bromide assay in table form. Screening the individual components of this mixture afforded the direct distamycin A analog **49** as having the highest affinity, followed closely by **53** containing the thiophene subunit at the first position (A). The same overall pattern was observed with poly[dA]–poly[dT], where **49** and **53** also showed the highest affinity (Table 3). Both these agents exhibited diminished affinity for the ARE-consensus sequence presumably resulting from the intervening GC base pair. Two additional mixtures, **109–118** and **119–128**, also bound the PSA-ARE-3 sequence effectively with the general trend **119–128** > **49–58** > **109–118**, the same general trend seen with poly[dA]–poly[dT] (Figures 15 and 16). The individual trimers **124** and **128** displayed tight binding to the PSA-ARE-3 sequence analogous to **49** and **53**. Importantly, **124** showed a loss of affinity to the ARE-consensus analogous to **49** and **53**, but **128** retained equal affinity making this agent ideal in maintaining high affinity for both the PSA-ARE-3 and ARE-consensus sequences.

Figure 31 shows the hairpin DNA oligomer used in the binding affinity assay. A survey of distamycin A binding to all possible 5 base pair DNA sequences was conducted using a library of 512 of these hairpin DNA oligonucleotides containing all possible five base pair sequences of the general format 5'-GCXXXXXC-3' with a 5-A loop. Although there are 1024 possible sequences containing 5 base pairs, two complementary sequences are contained in each hairpin differing only in their location relative to the position of the adenine loop making, for example, the sequence 5'-ATGCA equivalent to the sequence 5'-TGCAT as shown in the lower portion of the figure.

Figure 32 shows the results of screening the 512-membered library of

hairpin oligo-nucleotides using distamycin A. As expected, affinity increases with increasing AT content. The top sequences include the sites 5'-ATAA, 5'-AATT, 5'-AAAT, and 5'-AAAA and among the twenty hairpins showing the greatest decrease in % fluorescence, three four base-pair sequences occur most often: 5'-AATT, 5'-AAAT, 5'-AATA.

Figure 33 is a table showing the few absolute binding constants for distamycin A to short AT-rich sequences that have been published. The comparison of all those disclosed show the relative trend 5'-AATTT>AAAAA>AATAA>ATTAA (Rentzeperis, D.; et al. *Biochemistry* **1995**, *34*, 2937 and Wade, W. S.; Mrksich, M.; Dervan, P. B. *Biochemistry* **1993**, *32*, 11385). The ethidium bromide displacement assay revealed the same general trend and a quantitative titration measurement of binding constants with the hairpin oligo-nucleotides containing these sequences afforded binding constants that are not only consistent with the relative trend (Figure 21), but also within a factor of 2–3 of *all* the absolute binding constants previously determined through calorimetry and footprinting (Table 4). Given that the DNA upon which the measurements were made is different, that the buffer conditions are not identical, and that entries 2–4 were derived from a close analog of distamycin A, all which may contribute to small discrepancies in the absolute binding constants, the ethidium bromide displacement titration assay appears to be remarkably accurate at reproducing absolute binding constants.

Figure 34 shows the ethidium bromide binding constants obtained from Baguley, B.C.; Falkenhaug, E.-M. *Nucleic Acid Res.* **1978**, *5*, 161. The trend displayed is that the ethidium bromide binding constant varies considerably and the displacement does not follow a 1:1 stoichiometry. Both factors complicate the use of a competitive binding model for establishing binding constants.

Figure 35 is a bar chart showing the 20 highest affinity sequences. The sequence selectivity of compound **128** was established by screening it against the library of 512 hairpin oligonucleotides. It was found to clearly bind with a selectivity distinct from that of distamycin A and it appears to exhibit a significant preference for PuPyPy (purine-pyrimidine-pyrimidine) sequences. Of the 20 highest affinity sequences, 16 contain the PuPyPy motif (80%), where statistically 37.5% of the sequences would be expected to contain this motif in a random

sample. One of the four exceptions contained a five-base-pair AT-rich site. Within both of the androgen response elements used to identify **128**, the PuPyPy motif is repeated three times. It appears that this may be the reason for the equally high binding affinity of **128** with both sequences.

Figure 36 shows a number of derivatives of distamycin A where the side chains were altered in a systematic manner. Analysis of these derivatives using the quantitative titration with displacement of prebound ethidium bromide showed that there is very little difference between an amidine as the basic side chain and the dimethylamino group (**Distamycin** vs. **130**) or between placing the basic side chain at the *C*- or *N*-terminal end of the trimer (**49** vs. **129**). Amine **129** shows slightly lower binding affinity to poly[dA]–poly[dT] than **49** which may arise from the incorporation of a bulky *t*-butyl group present at the *N*-terminus. Interestingly, there is approximately a three-fold difference in binding between distamycin A or **130** and the tripyrrole **49**. The primary difference between the two molecules is the presence of an additional potential hydrogen bond donor in distamycin and **130** (*N*-terminal formamide) that is not present in either **49** (*C*-terminal ester), **129** (*N*-terminal BOC-group), or **132** (*C*-terminal dimethylamide). When a potential hydrogen bond donor group is included at the *C*-terminus (**131**), the binding affinity does approximate that of distamycin A. The difference in free energy of binding between those molecules containing an additional donor hydrogen bonding group (distamycin A and **130–131**) and those which do not (**49**, **129**, **132**) is approximately 1 kcal/mol, the value of a single hydrogen bond. Interestingly, adding a second substituent containing an additional basic, protonated amine (**133** vs **131**) does not further increase the DNA binding affinity.

Figure 37 shows a number of derivatives of the trimer core of **128** where the side chains were altered in a systematic manner. Analysis of these derivatives using the quantitative titration with displacement of prebound ethidium bromide showed that there is very little difference between an amidine as the basic side chain and the dimethylamino group (**136** vs. **138**) or between placing the basic side chain at the *C*- or *N*-terminal end of the trimer (**138** vs. **128**). Amine **128** shows slightly lower binding affinity to poly[dA]–poly[dT] than **138**. Interestingly, there is approximately a three-fold difference in binding between **136** and **138**. The primary difference between the two molecules is the presence

of an additional potential hydrogen bond donor in **136** (*N*-terminal formamide) that is not present in either **128** (*C*-terminal ester), **138** (*N*-terminal BOC-group), or **135** (*C*-terminal dimethylamide). The difference in free energy of binding between those molecules containing an additional donor hydrogen bonding group (**136**, **134**) and those which do not (**138**, **135**) is approximately 0.7 kcal/mol, which is 70% of the value of a single hydrogen bond. Adding a second substituent containing an additional basic, protonated amine (**128** vs **137**) does not further increase the DNA binding affinity.

What is claimed is:

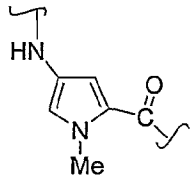
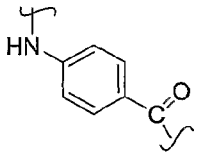
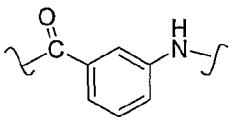
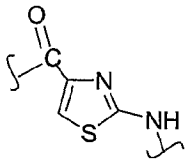
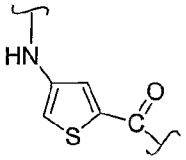
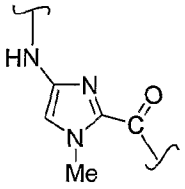
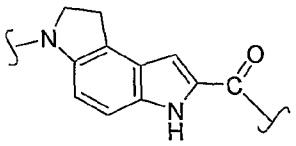
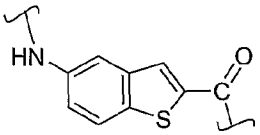
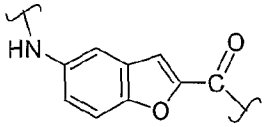
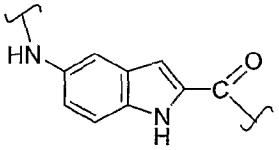
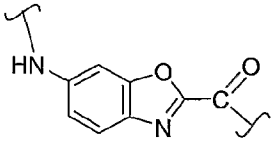
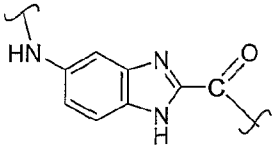
1. An analog of distamycin A represented by the following structure:

wherein:

R is a radical selected from the group consisting of -C(O)O-(C1-C6 alkyl) and $-C(O)CH_2CH_2CH_2NMe_2$;

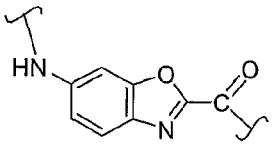
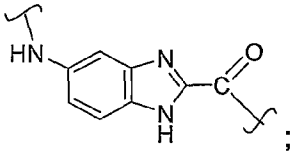
R' is -O(C1-C6 alkyl); and

-NH-Subunit **A**-C(O)- , -NH-Subunit **B**-C(O)- , and -NH-Subunit **C**-C(O)- are each a diradical independently selected from the group consisting of the following structures:

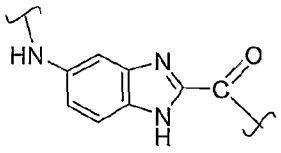













with the following provisos:

-NH-Subunit **A**-C(O)- can not be represented by either of the following structures:

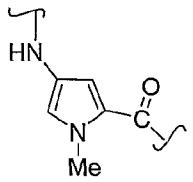

or

;

-NH-Subunit **B**-C(O)- can not be represented by following structure;



; and

-NH-Subunit **A**-C(O)- , -NH-Subunit **B**-C(O)- , and -NH-Subunit **C**-C(O)- can not all simultaneously be represented by the following structure:



.

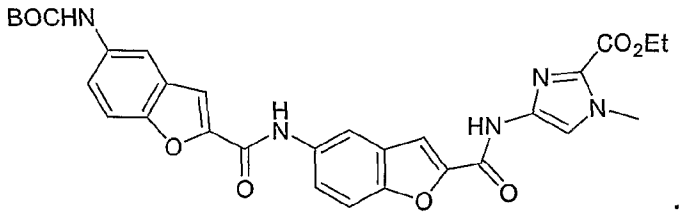
2. An analog of distamycin A according to Claim 1 wherein **R** is -C(O)O-*t*Bu.

3. An analog of distamycin A according to Claim 1 wherein **R** is -$C(O)CH_2CH_2CH_2NMe_2$.

4. An analog of distamycin A according to Claim 1 wherein **R′** is selected from the group consisting of -OMe and -OEt.

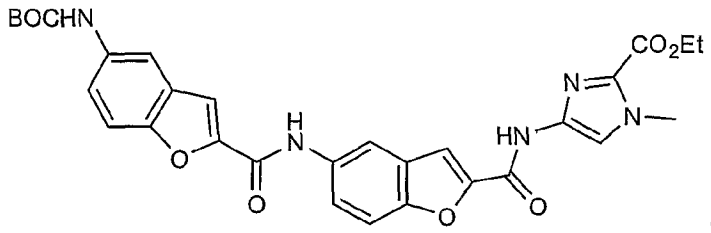
5. An analog of distamycin A according to Claim 1 wherein there is a proviso that
-NH-Subunit **A**-C(O)- , -NH-Subunit **B**-C(O)- , and -NH-Subunit **C**-C(O)- can not all be identical.

6. An analog of distamycin A according to Claim 5 represented by the following structure:



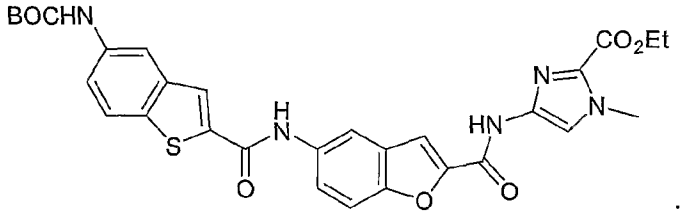
.

7. An analog of distamycin A according to Claim 5 represented by the following structure:

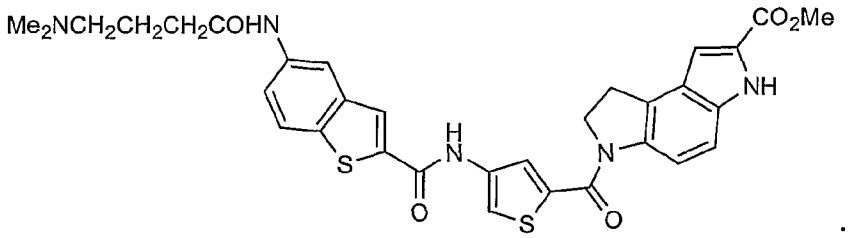


8. An analog of distamycin A according to Claim 1 wherein there is a proviso that none of -NH-Subunit **A**-C(O)- , -NH-Subunit **B**-C(O)- , and -NH-Subunit **C**-C(O)- are identical.

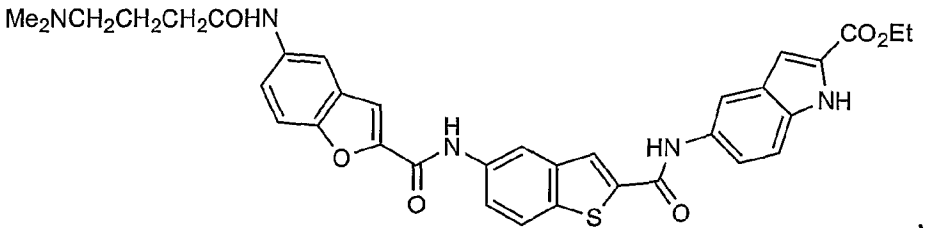
9. An analog of distamycin A according to Claim 8 represented by the following structure:



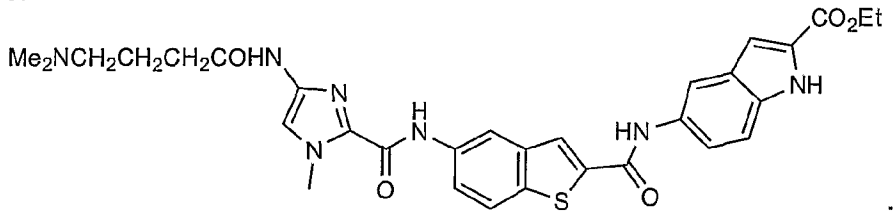
10. An analog of distamycin A according to Claim 8 represented by the following structure:



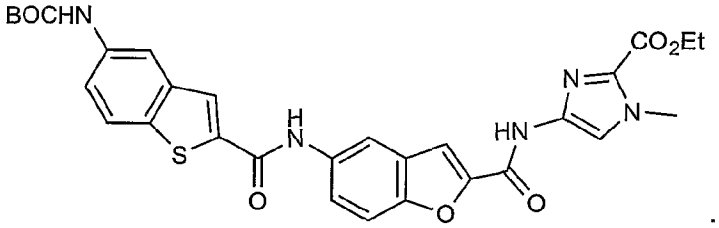
11. An analog of distamycin A according to Claim 8 represented by the following structure:



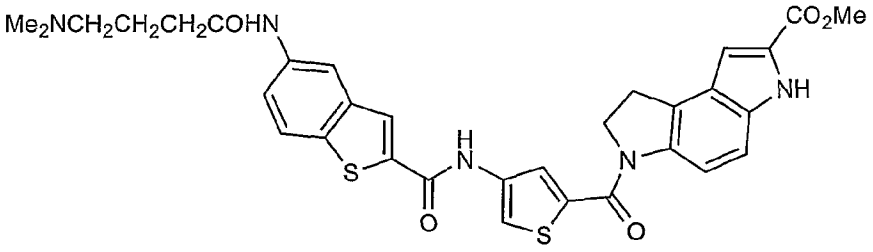
12. An analog of distamycin A according to Claim 8 represented by the following structure:



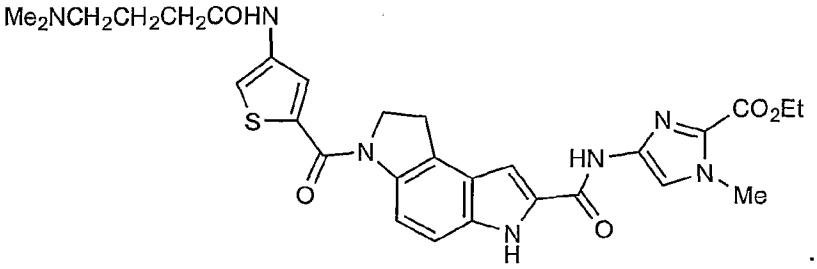
13. An analog of distamycin A according to Claim 8 represented by the following structure:



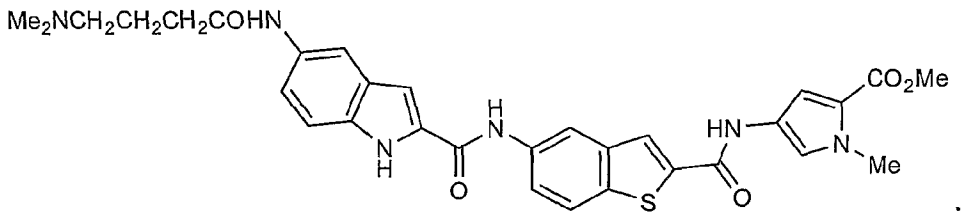
14. An analog of distamycin A according to Claim 8 represented by the following structure:



15. An analog of distamycin A according to Claim 8 represented by the following structure:



16. An analog of distamycin A according to Claim 8 represented by the following structure:



17. A positional scanning library comprising a collection of ten or more of the compounds of claims 1-16.

18. A process for synthesizing a library of amide linked aromatic trimers represented by the following structure:

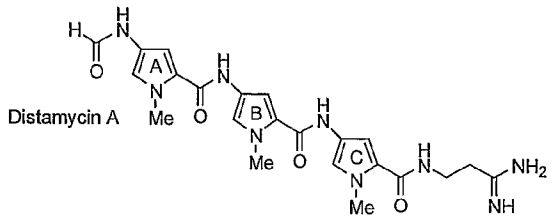
wherein Subunit **A** is any aromatic radical of a plurality of aromatic radicals, Subunit **B** is a first aromatic radical, and Subunit **C** is a second aromatic radical, the process comprising the following steps:

Step A: linking Subunit **B** to Subunit **C** by means of a first amide linkage to form a dimer of the first and second aromatic radicals, the dimer being represented by the following structure:

; and then

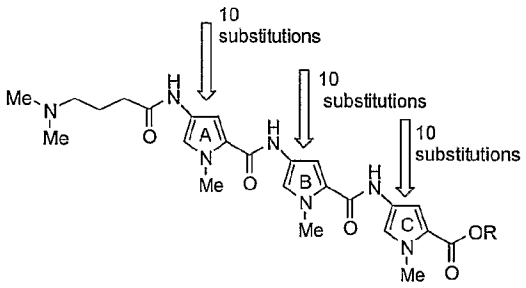
Step B: linking a plurality of the dimers of said Step A to a plurality of Subunits **A** by means of a second amide linkage for forming said library of compounds, each element of said library being a trimer of aromatic radicals linked by amide linkages.

19. A process for killing a cancer cell comprising the step of contacting the cancer cell with a solution containing a cytotoxic concentration of a compound of Claims 1-16.



Solution phase combinatorial chemistry using different heterocyclic amino acids and liquid–liquid purification protocols

Substitute with different heterocyclic amino acids



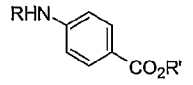
1000-member library of distamycin A analogues

Sublibraries I	Sublibraries II	Sublibraries III
BOC–**A**$_1$–X–X–OR	BOC–X–**B**$_1$–X–OR	BOC–X–X–**C**$_1$–OMe
BOC–**A**$_2$–X–X–OR	BOC–X–**B**$_2$–X–OR	BOC–X–X–**C**$_2$–OMe
BOC–**A**$_3$–X–X–OR	BOC–X–**B**$_3$–X–OR	BOC–X–X–**C**$_3$–OEt
⋮	⋮	⋮
BOC–**A**$_{10}$–X–X–OR	BOC–X–**B**$_{10}$–X–OR	BOC–X–X–**C**$_{10}$–OEt

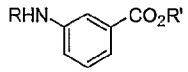
X = variable position containing full mixture of 10 monomers

bold letter = fixed position within the tripeptide containing one of the 10 monomers

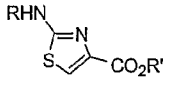
Figure 1



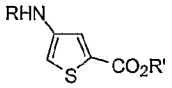
5a R = H, R' = Me
5b R = $CO_2{}^t$Bu, R' = H



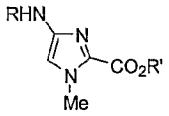
6a R = H, R' = Me
6b R = $CO_2{}^t$Bu, R' = H



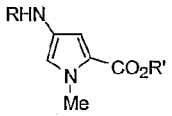
7a R = H, R' = Et
7b R = $CO_2{}^t$Bu, R' = H



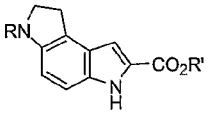
8a R = H, R' = Me
8b R = $CO_2{}^t$Bu, R' = H



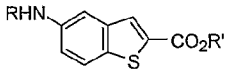
9a R = H, R' = Et
9b R = $CO_2{}^t$Bu, R' = H



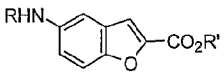
1a R = H, R' = Me
1b R = $CO_2{}^t$Bu, R' = H



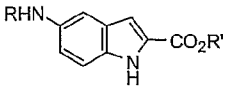
10a R = H, R' = Me
10b R = $CO_2{}^t$Bu, R' = H



11a R = H, R' = Me
11b R = $CO_2{}^t$Bu, R' = H



12a R = H, R' = Me
12b R = $CO_2{}^t$Bu, R' = H



13a.R = H, R' = Et
13b R = $CO_2{}^t$Bu, R' = H

Figure 2

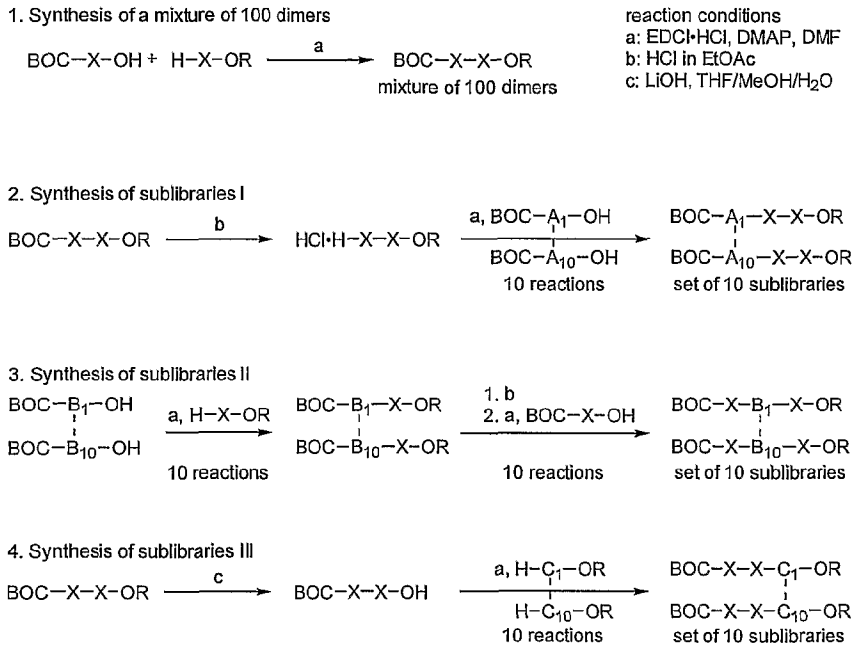

1. Synthesis of a mixture of 100 dimers
BOC–X–OH + H–X–OR
a
BOC–X–X–OR
mixture of 100 dimers
reaction conditions
a: EDCI•HCl, DMAP, DMF
b: HCl in EtOAc
c: LiOH, THF/MeOH/H2O
2. Synthesis of sublibraries I
BOC–X–X–OR
b
HCl•H–X–X–OR
a, BOC–A1–OH
BOC–A10–OH
10 reactions
BOC–A1–X–X–OR
BOC–A10–X–X–OR
set of 10 sublibraries
3. Synthesis of sublibraries II
BOC–B1–OH
BOC–B10–OH
a, H–X–OR
10 reactions
BOC–B1–X–OR
BOC–B10–X–OR
1. b
2. a, BOC–X–OH
10 reactions
BOC–X–B1–X–OR
BOC–X–B10–X–OR
set of 10 sublibraries
4. Synthesis of sublibraries III
BOC–X–X–OR
c
BOC–X–X–OH
a, H–C1–OR
H–C10–OR
10 reactions
BOC–X–X–C1–OR
BOC–X–X–C10–OR
set of 10 sublibraries

Figure 3

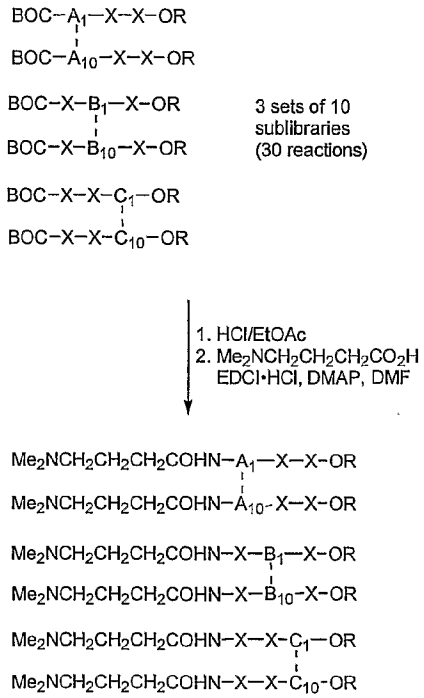

BOC–A1–X–X–OR
BOC–A10–X–X–OR
BOC–X–B1–X–OR
BOC–X–B10–X–OR
BOC–X–X–C1–OR
BOC–X–X–C10–OR
3 sets of 10 sublibraries (30 reactions)
1. HCl/EtOAc
2. Me2NCH2CH2CH2CO2H EDCI•HCl, DMAP, DMF
Me2NCH2CH2CH2COHN–A1–X–X–OR
Me2NCH2CH2CH2COHN–A10–X–X–OR
Me2NCH2CH2CH2COHN–X–B1–X–OR
Me2NCH2CH2CH2COHN–X–B10–X–OR
Me2NCH2CH2CH2COHN–X–X–C1–OR
Me2NCH2CH2CH2COHN–X–X–C10–OR

Figure 4

Yields of BOC- and DMABA-trimer library synthesis

subunit number	Yield (%) BOC-trimers			Yield (%) DMABA-trimers		
	A_{1-10}	B_{1-10}	C_{1-10}	A_{1-10}	B_{1-10}	C_{1-10}
1	77	77	45	54	12	42
2	72	72	71	32	54	54
3	68	68	40	43	22	47
4	77	77	70	50	33	57
5	92	92	51	39	29	64
6	79	79	73	13	62	76
7	99	99	81	26	41	65
8	76	76	58	53	45	46
9	71	71	58	55	39	65
10	69	69	64	49	38	79

Figure 5

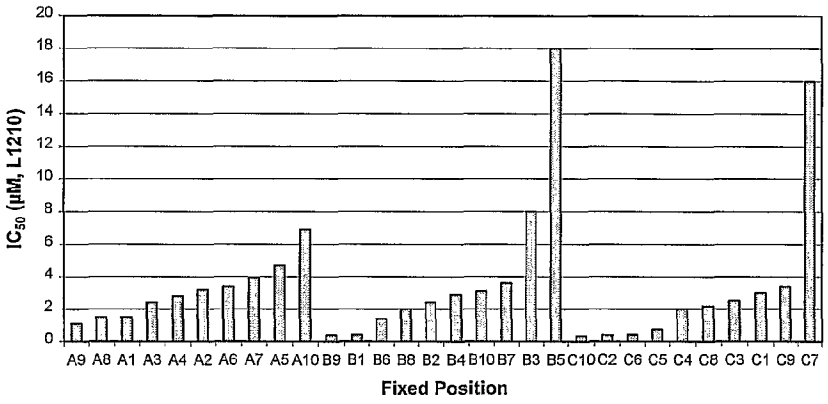
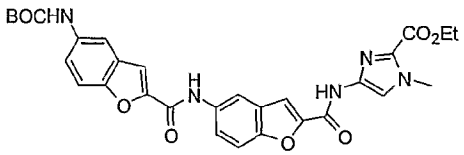

20
18
16
14
12
10
8
6
4
2
0
IC50 (μM, L1210)
A9 A8 A1 A3 A4 A2 A6 A7 A5 A10 B9 B1 B6 B8 B2 B4 B10 B7 B3 B5 C10 C2 C6 C5 C4 C8 C3 C1 C9 C7
Fixed Position

Figure 6

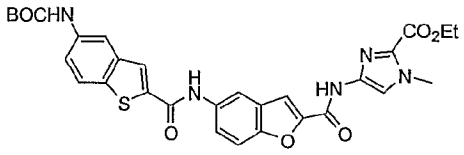
Cytotoxic activity of candidate compounds

Compound	IC_{50} (μM, L1210)
BOC-A_9-B_9-C_{10}-OEt	1.4
BOC-A_9-B_9-C_2-OMe	>100
BOC-A_9-B_9-C_6-OMe	1.8
BOC-A_9-B_9-C_5-OEt (**67**)	0.029
BOC-A_9-B_1-C_{10}-OEt	2.7
BOC-A_9-B_1-C_2-OMe	> 100
BOC-A_9-B_1-C_6-OMe	> 100
BOC-A_9-B_1-C_5-OEt (**200**)	0.42
BOC-A_8-B_9-C_{10}-OEt	3.4
BOC-A_8-B_9-C_2-OMe	51
BOC-A_8-B_9-C_6-OMe	46
BOC-A_8-B_9-C_5-OEt (**66**)	0.069
BOC-A_8-B_1-C_{10}-OMe	5.2
BOC-A_8-B_1-C_2-OMe	> 100
BOC-A_8-B_1-C_6-OMe	> 100
BOC-A_8-B_1-C_5-OEt	3.7

Figure 7



67



66

Figure 8

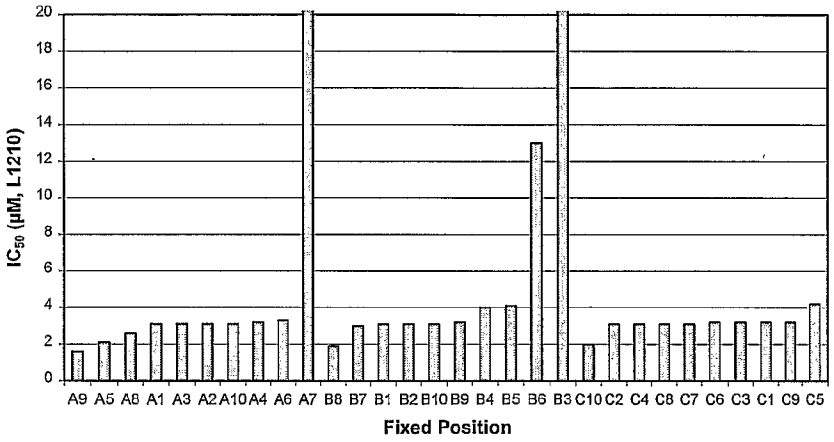
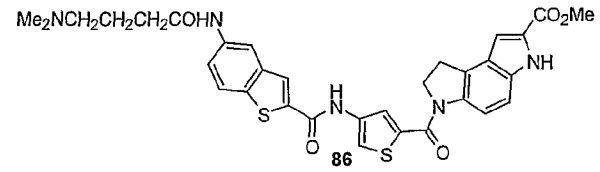
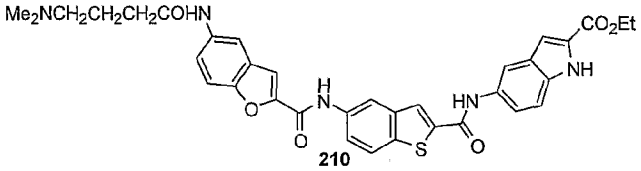

20
18
16
14
12
10
8
6
4
2
0
IC50 (μM, L1210)
A9 A5 A8 A1 A3 A2 A10 A4 A6 A7 B8 B7 B1 B2 B10 B9 B4 B5 B6 B3 C10 C2 C4 C8 C7 C6 C3 C1 C9 C5
Fixed Position

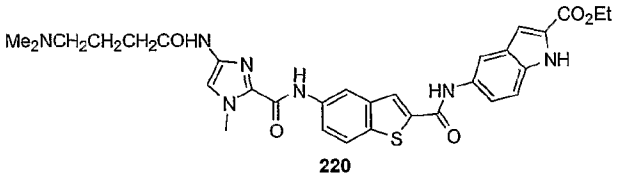
Figure 9



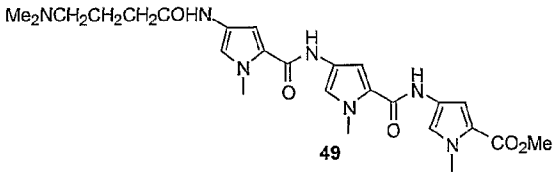
86



210



220



49

Figure 10

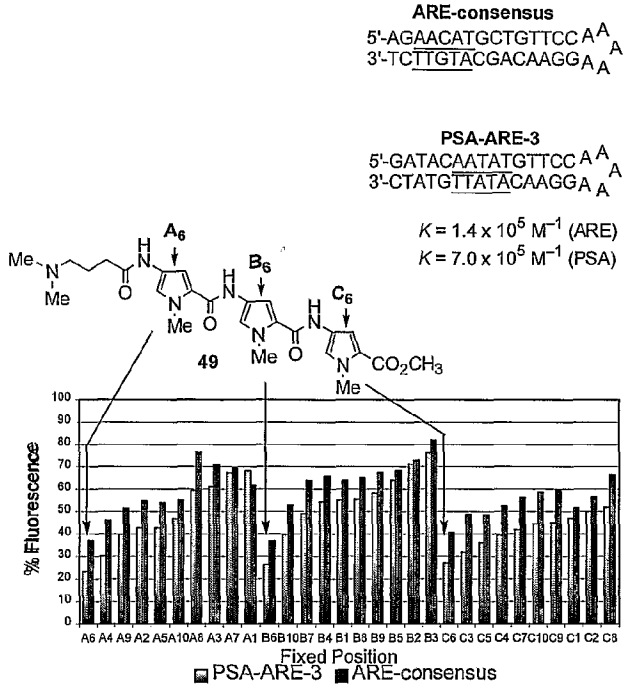
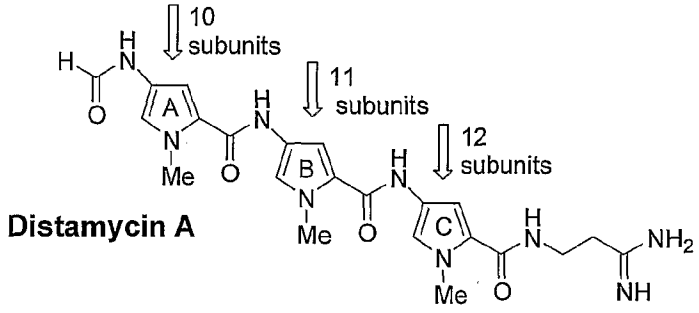

ARE-consensus
5'-AGAACATGCTGTTCC A A A
3'-TCTTGTACGACAAGG A A A
PSA-ARE-3
5'-GATACAATATGTTCC A A A
3'-CTATGTTATACAAGG A A A
K = 1.4 x 10^5 M^-1 (ARE)
K = 7.0 x 10^5 M^-1 (PSA)
A6
B6
C6
49
% Fluorescence
Fixed Position
PSA-ARE-3
ARE-consensus

Figure 11

Substitute with different heterocyclic amino acids



Distamycin A

Solution phase combinatorial chemistry using different heterocyclic amino acids and liquid–liquid purification protocols

First generation libraries of potential DNA binding agents

Derivatize with a basic side chain

Second generation libraries of potential DNA binding agents with increased affinity

Figure 12

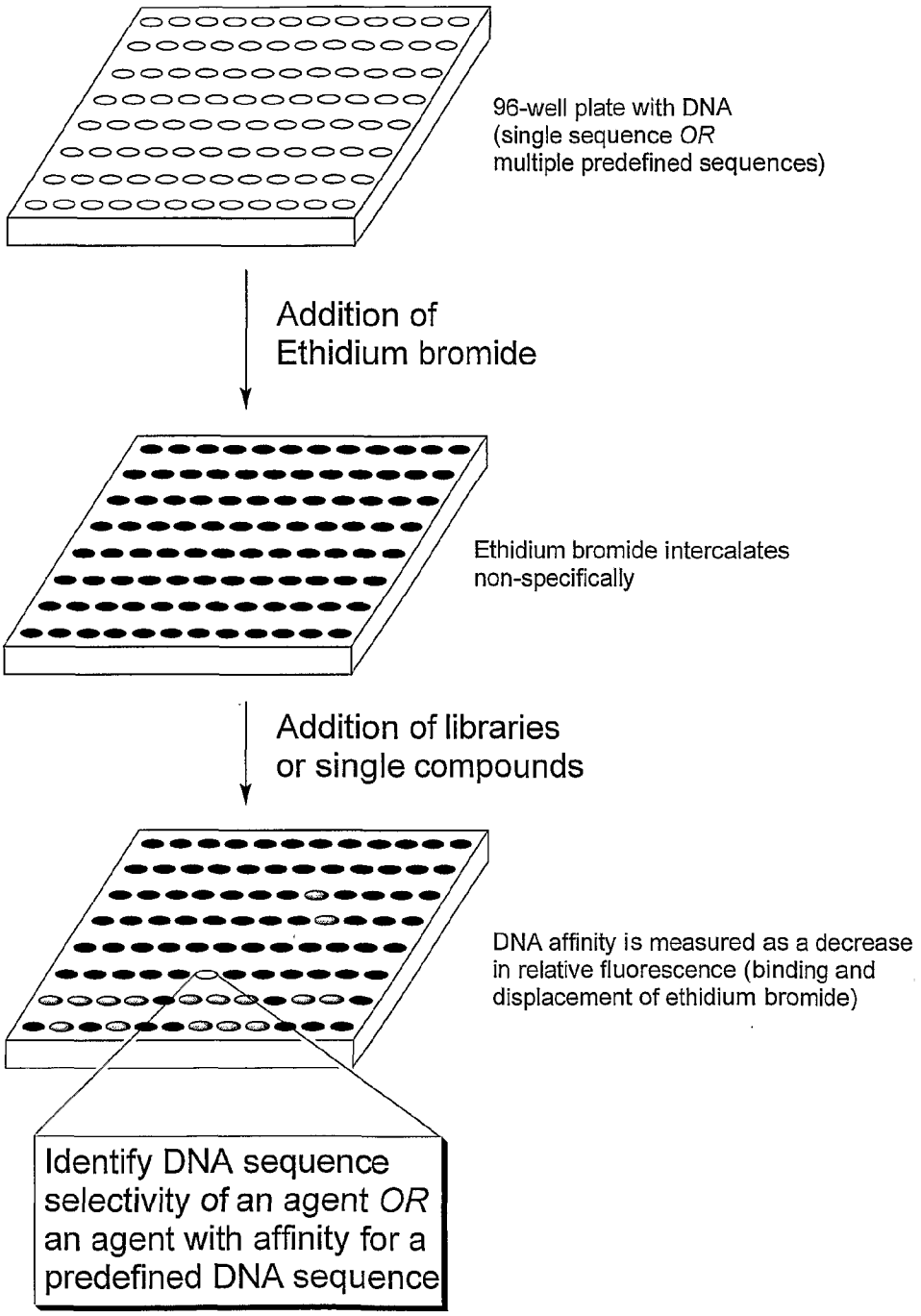

96-well plate with DNA
(single sequence OR
multiple predefined sequences)
Addition of
Ethidium bromide
Ethidium bromide intercalates
non-specifically
Addition of libraries
or single compounds
DNA affinity is measured as a decrease
in relative fluorescence (binding and
displacement of ethidium bromide)
Identify DNA sequence
selectivity of an agent OR
an agent with affinity for a
predefined DNA sequence

Figure 13

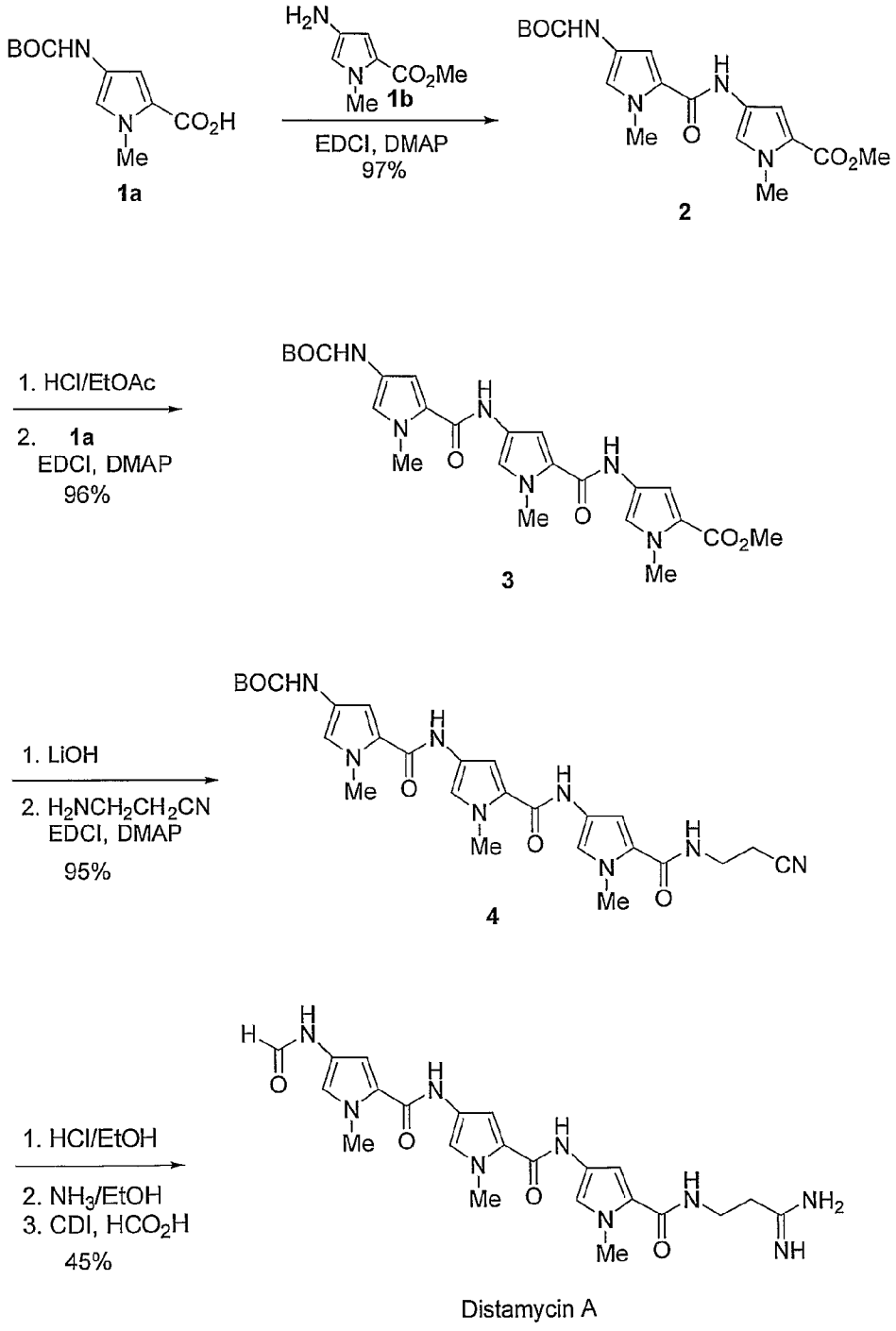

BOCHN
CO2H
N
Me
1a
H2N
CO2Me
N
Me 1b
EDCI, DMAP
97%
BOCHN
H
N
N
Me
O
CO2Me
N
Me
2
1. HCl/EtOAc
2. 1a
EDCI, DMAP
96%
BOCHN
3
1. LiOH
2. H2NCH2CH2CN
EDCI, DMAP
95%
CN
4
1. HCl/EtOH
2. NH3/EtOH
3. CDI, HCO2H
45%
NH2
NH
Distamycin A

Figure 14

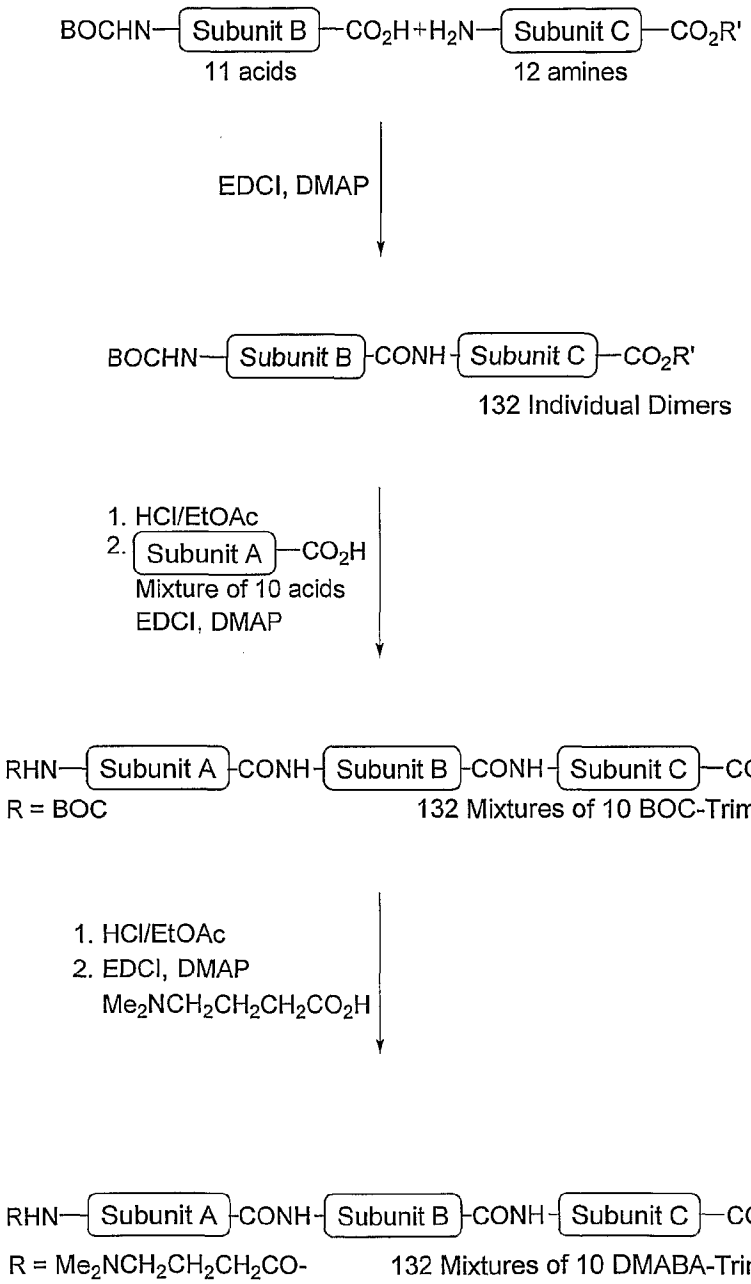
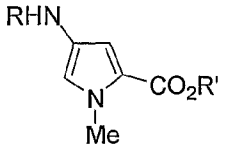
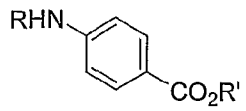

BOCHN—Subunit B—CO2H+H2N—Subunit C—CO2R'
11 acids
12 amines
EDCI, DMAP
BOCHN—Subunit B—CONH—Subunit C—CO2R'
132 Individual Dimers
1. HCl/EtOAc
2. Subunit A—CO2H
Mixture of 10 acids
EDCI, DMAP
RHN—Subunit A—CONH—Subunit B—CONH—Subunit C—CO2R'
R = BOC
132 Mixtures of 10 BOC-Trimers
1. HCl/EtOAc
2. EDCI, DMAP
Me2NCH2CH2CH2CO2H
RHN—Subunit A—CONH—Subunit B—CONH—Subunit C—CO2R'
R = Me2NCH2CH2CH2CO-
132 Mixtures of 10 DMABA-Trimers

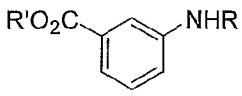
Figure 15



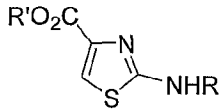
1a R = H, R' = Me
1b R = $CO_2{}^tBu$, R' = H



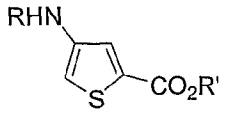
5a R = H, R' = Me
5b R = $CO_2{}^tBu$, R' = H



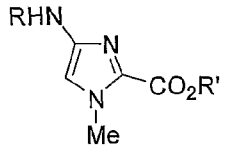
6a R = H, R' = Me
6b R = $CO_2{}^tBu$, R' = H



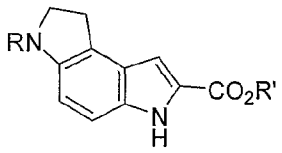
7a R = H, R' = Et
7b R = $CO_2{}^tBu$, R' = H



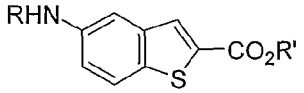
8a R = H, R' = Me
8b R = $CO_2{}^tBu$, R' = H



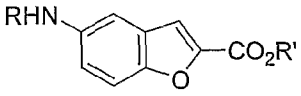
9a R = H, R' = Et
9b R = $CO_2{}^tBu$, R' = H



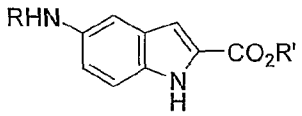
10a R = H, R' = Me
10b R = $CO_2{}^tBu$, R' = H



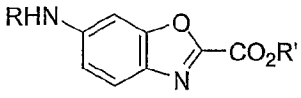
11a R = H, R' = Me
11b R = $CO_2{}^tBu$, R' = H



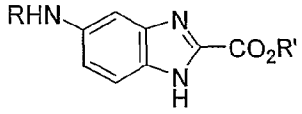
12a R = H, R' = Me
12b R = $CO_2{}^tBu$, R' = H



13a R = H, R' = Et
13b R = $CO_2{}^tBu$, R' = H



14a R = H, R' = Me
14b R = $CO_2{}^tBu$, R' = Na



15a R = H, R' = Me
15b R = $CO_2{}^tBu$, R' = Na

Figure 16

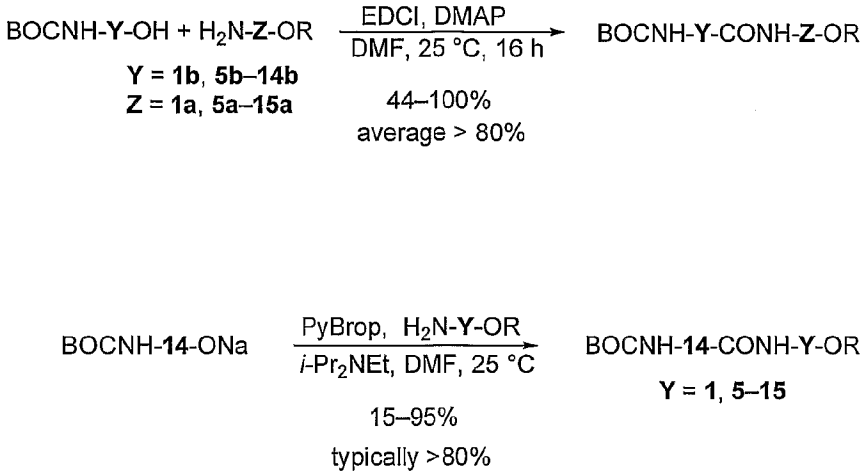

BOCNH-Y-OH + H2N-Z-OR
Y = 1b, 5b–14b
Z = 1a, 5a–15a
EDCI, DMAP
DMF, 25 °C, 16 h
44–100%
average > 80%
BOCNH-Y-CONH-Z-OR
BOCNH-14-ONa
PyBrop, H2N-Y-OR
i-Pr2NEt, DMF, 25 °C
15–95%
typically >80%
BOCNH-14-CONH-Y-OR
Y = 1, 5–15

Figure 17

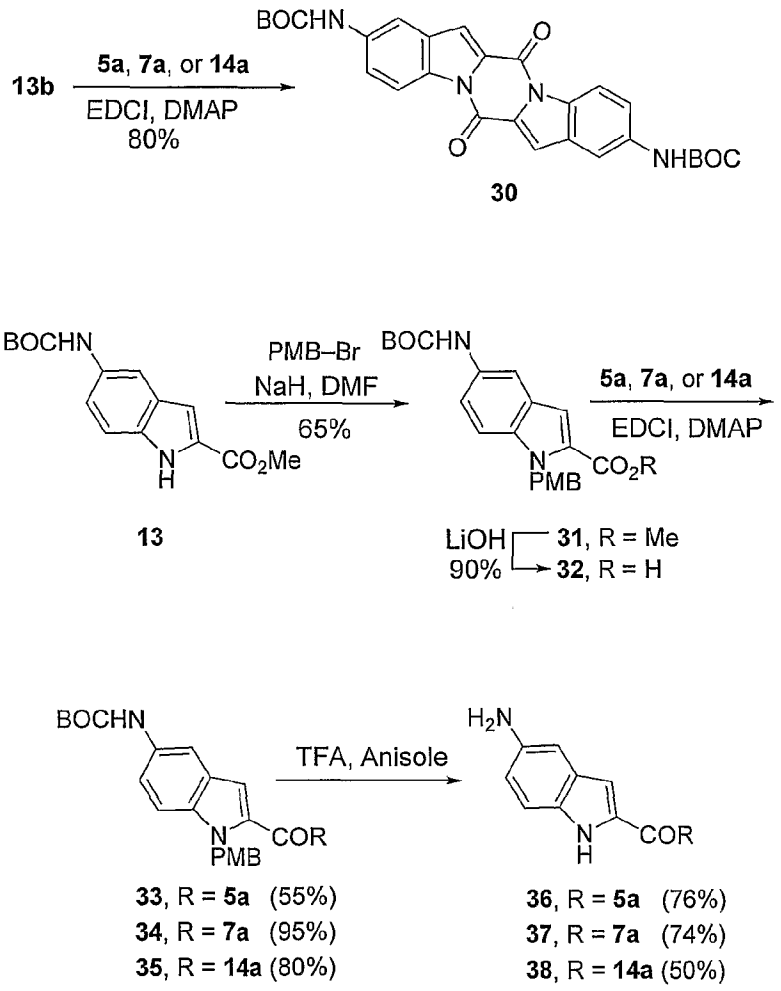
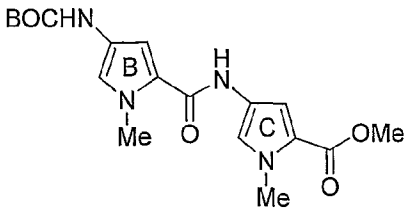

BOCHN
13b
5a, 7a, or 14a
EDCI, DMAP
80%
NHBOC
30
PMB–Br
NaH, DMF
65%
CO2Me
13
LiOH
90%
31, R = Me
32, R = H
EDCI, DMAP
CO2R
PMB
TFA, Anisole
H2N
COR
33, R = 5a (55%)
34, R = 7a (95%)
35, R = 14a (80%)
36, R = 5a (76%)
37, R = 7a (74%)
38, R = 14a (50%)

Figure 18



BOCHN-**1**-COHN-**1**-OMe

X = **1b**, **5**–**13b** | 1. HCl/EtOAc
2. BOCHN-**X**-OH,
EDCI, DMAP, DMF

BOCHN-**X**-COHN-**1**-COHN-**1**-OMe

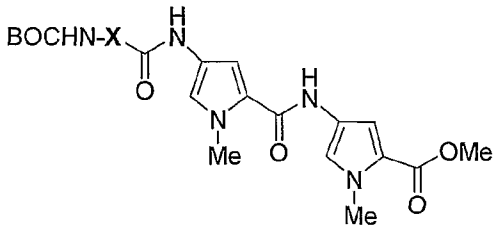
cmpd	X	% Yield
39	**1**	92
40	**5**	95
41	**6**	91
42	**7**	95
43	**8**	95
44	**9**	60
45	**10**	92
46	**11**	88
47	**12**	76
48	**13**	95

BOCHN-**Y**-CONH-**Z**-OR

Y = **1**, **5**–**14**
Z = **1**, **5**–**15** | 1. HCl/EtOAc
2. BOCHN-**X**-OH
EDCI, DMAP, DMF

BOCNH-**X**-CONH-**Y**-CONH-**Z**-OR
X = mixture of **1**, **5**–**13**
26–100 %
Average > 60%

Figure 19



BOCHN-**X**-COHN-**1**-COHN-**1**-OMe

1. HCl/ EtOAc
2. $Me_2NCH_2CH_2CH_2CO_2H$, EDCI, DMAP, DMF

$Me_2NCH_2CH_2CH_2$COHN-**X**-COHN-**1**-COHN-**1**-OMe

cmpd	X	% Yield
49	**1**	73
50	**5**	50
51	**6**	61
52	**7**	74
53	**8**	48
54	**9**	87
55	**10**	35
56	**11**	61
57	**12**	71
58	**13**	23

BOCHN-**X**-COHN-**Y**-COHN-**Z**-OR

X = mixture of **1**, **5–13**
Y = **1**, **5–14**
Z = **1**, **5–15**

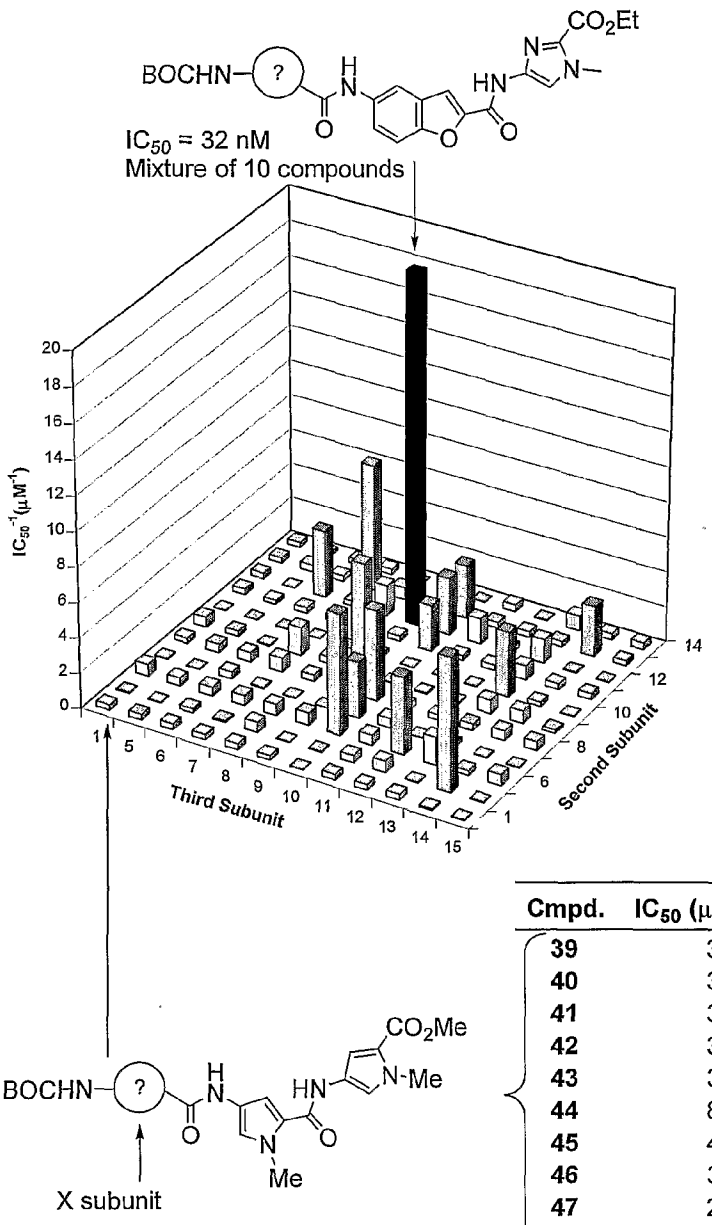
1. HCl/ EtOAc
2. $Me_2NCH_2CH_2CH_2CO_2H$, EDCI, DMAP, DMF

$Me_2NCH_2CH_2CH_2CO_2$HN-**X**-COHN-**Y**-COHN-**Z**-OR

22–100%

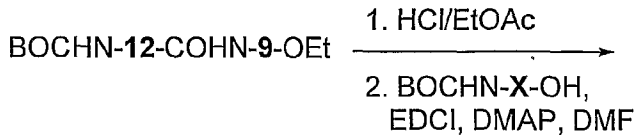
average > 60%

Figure 20



Cmpd.	IC_{50} (μM, L1210)
39	32
40	31
41	32
42	30
43	37
44	84
45	4.2
46	3.6
47	22
48	2.1
Mixture	3.3

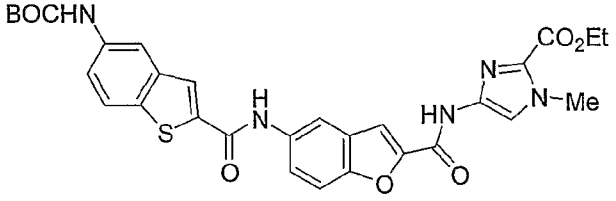
Figure 21



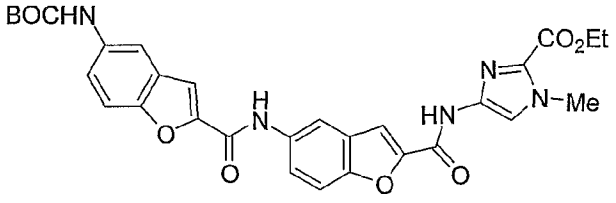
BOCHN-**X**-COHN-**12**-COHN-**9**-OEt

Table 1

cmpd	X	% Yield	IC_{50} (μM, L1210)
59	**1**	42	0.30
60	**5**	52	2.7
61	**6**	68	>100
62	**7**	54	0.34
63	**8**	89	1.5
64	**9**	65	1.8
65	**10**	55	1.3
66	**11**	63	0.069
67	**12**	69	0.029
68	**13**	72	1.4
Mixture			0.032

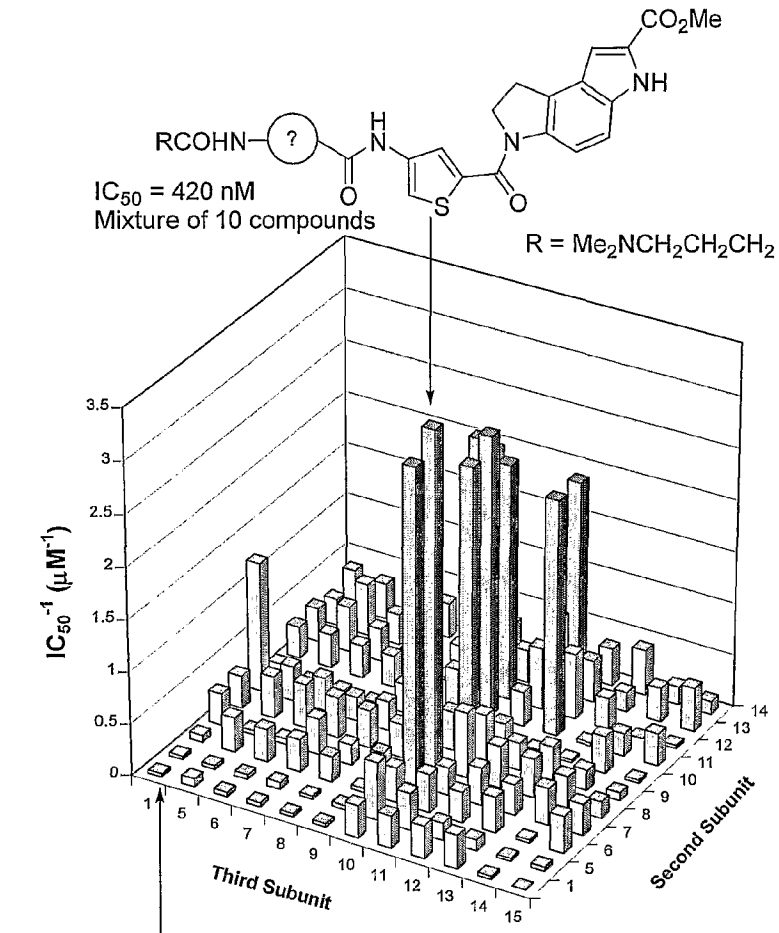


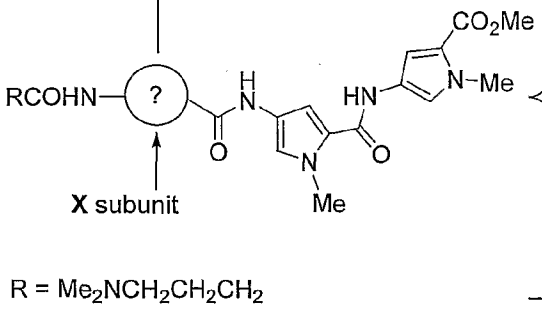
66



67

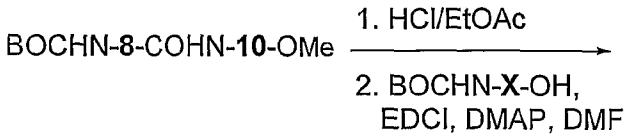
Figure 22

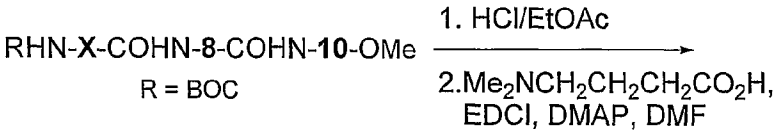




Cmpd.	IC_{50} (μM, L1210)
49	44
50	53
51	41
52	>100
53	>100
54	>100
55	51
56	46
57	33
58	33
Mixture	32

Figure 23

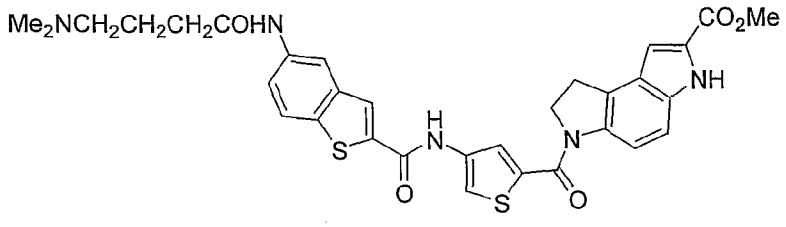




RHN-**X**-COHN-**8**-COHN-**10**-OMe

R = $Me_2NCH_2CH_2CH_2CO$

X	Yield (R = Boc)	IC_{50} (μM, L1210)	Yield (R = DMABA)	IC_{50} (μM, L1210)
1	**69**, 48%	0.82	**79**, 98%	220
5	**70**, 48%	25	**80**, 80%	230
6	**71**, 34%	2.3	**81**, 99%	420
7	**72**, 39%	1.9	**82**, 99%	200
8	**73**, 40%	7.5	**83**, 49%	20
9	**74**, 40%	45	**84**, 79%	64
10	**75**, 51%	0.80	**85**, 95%	>1000
11	**76**, 49%	35	**86**, 95%	0.46
12	**77**, 54%	67	**87**, 90%	7.6
13	**78**, 24%	27	**88**, 88%	>1000
Mixture		2.7		0.33



86

Figure 24

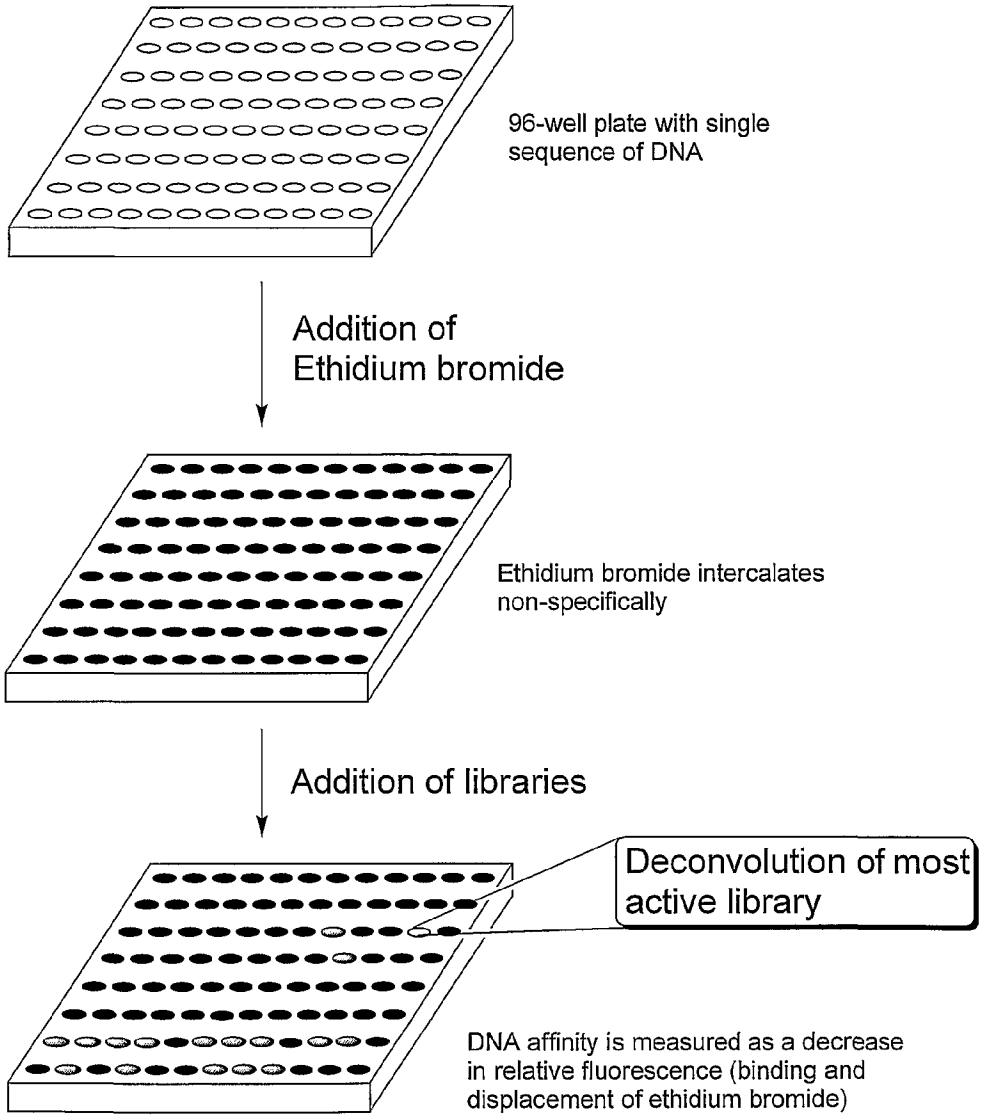
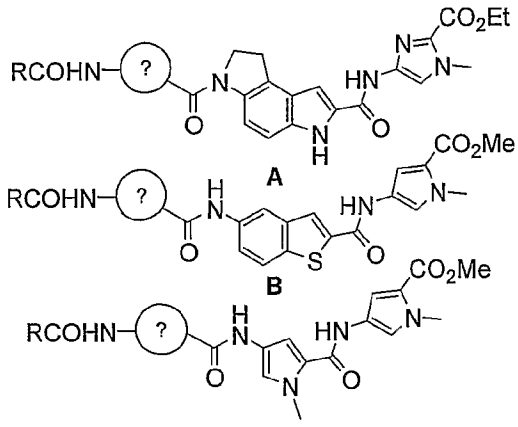
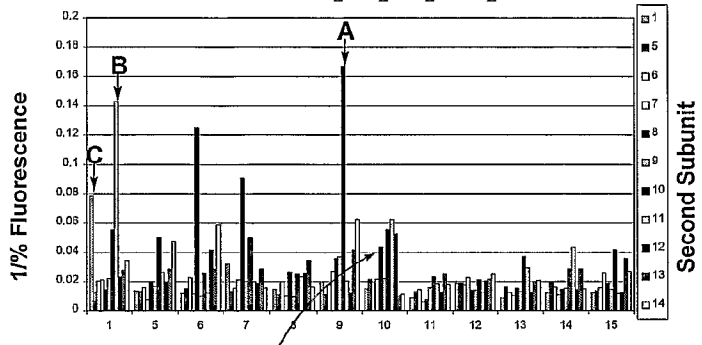

96-well plate with single sequence of DNA
Addition of Ethidium bromide
Ethidium bromide intercalates non-specifically
Addition of libraries
Deconvolution of most active library
DNA affinity is measured as a decrease in relative fluorescence (binding and displacement of ethidium bromide)

Figure 25



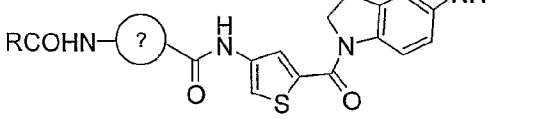
R = $CH_2CH_2CH_2NMe_2$

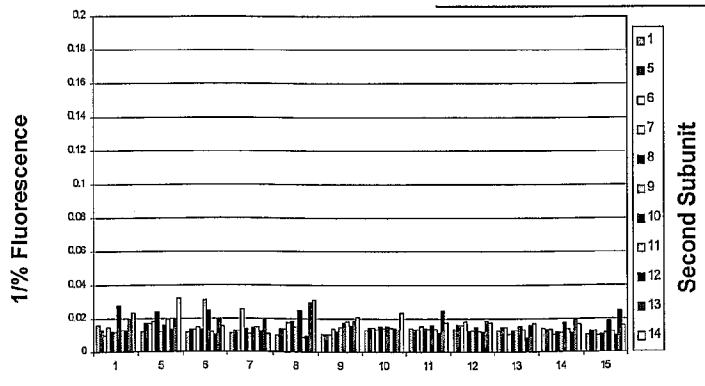


A

Third Subunit

Cmpd.	% Fluorescence (10μM)
79	34
80	100
81	58
82	72
83	77
84	67
85	72
86	54
87	90
88	81

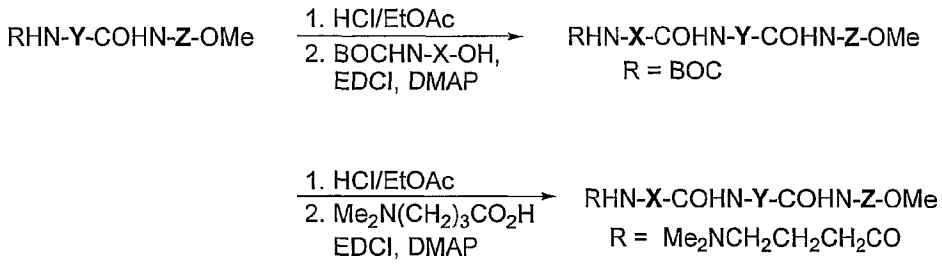




B

Third Subunit

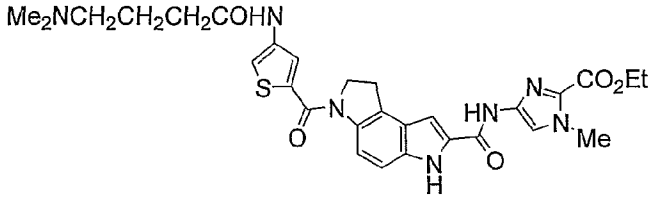
Figure 26



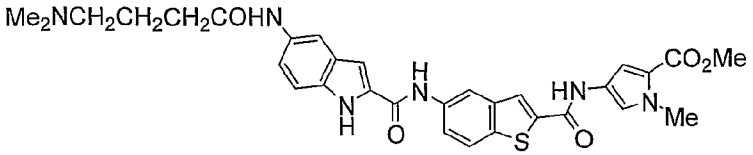
X	% Yield (R = BOC)		% Yield (R = $Me_2NCH_2CH_2CH_2CO$)	
	Y = 10, Z = 9	Y = 11, Z = 1	Y = 10, Z = 9	Y = 11, Z = 1
1	**89**, 60	**99**, 74	**109**, 78	**119**, 46
5	**90**, 66	**100**, 81	**110**, 67	**120**, 22
6	**91**, 63	**101**, 77	**111**, 99	**121**, 79
7	**92**, 82	**102**, 96	**112**, 61	**122**, 45
8	**93**, 50	**103**, 81	**113**, 83	**123**, 38
9	**94**, 46	**104**, 78	**114**, 78	**124**, 12
10	**95**, 83	**105**, 95	**115**, 50	**125**, 64
11	**96**, 90	**106**, 80	**116**, 74	**126**, 18
12	**97**, 82	**107**, 78	**117**, 68	**127**, 36
13	**98**, 94	**108**, 68	**118**, 42	**128**, 78

Figure 27

	IC_{50} (L1210, μM)			
1	**89**, 26	**99**, >100	**109**, >100	**119**, 32
5	**90**, 30	**100**, 1.7	**110**, 32	**120**, 4.7
6	**91**, 35	**101**, 0.13	**111**, 32	**121**, 32
7	**92**, 33	**102**, 0.24	**112**, 34	**122**, 3.3
8	**93**, 17	**103**, 0.65	**113**, 25	**123**, 32
9	**94**, 9.5	**104**, 6.3	**114**, 48	**124**, 32
10	**95**, 0.44	**105**, 0.55	**115**, 32	**125**, >100
11	**96**, 17	**106**, 32	**116**, 50	**126**, 33
12	**97**, 1	**107**, 1.3	**117**, 7.8	**127**, 3.2
13	**98**, 37	**108**, 2.6	**118**, 33	**128**, 32
Mixture	5.0	3.1	32	3.3
	% Fluorescence at 10 μM (Poly[dA]–Poly[dT])			
	49 29	**109** 74	**119** 61	
	50 100	**110** 79	**120** 78	
	51 70	**111** 73	**121** 34	
	52 71	**112** 33	**122** 66	
	53 25	**113** 67	**123** 83	
	54 78	**114** 46	**124** 49	
	55 72	**115** 74	**125** 56	
	56 53	**116** 51	**126** 91	
	57 67	**117** 37	**127** 72	
	58 54	**118** 35	**128** 9	



112, $K = 2.5 \times 10^6\ M^{-1}$



128, $K = 5.6 \times 10^6\ M^{-1}$

Figure 28

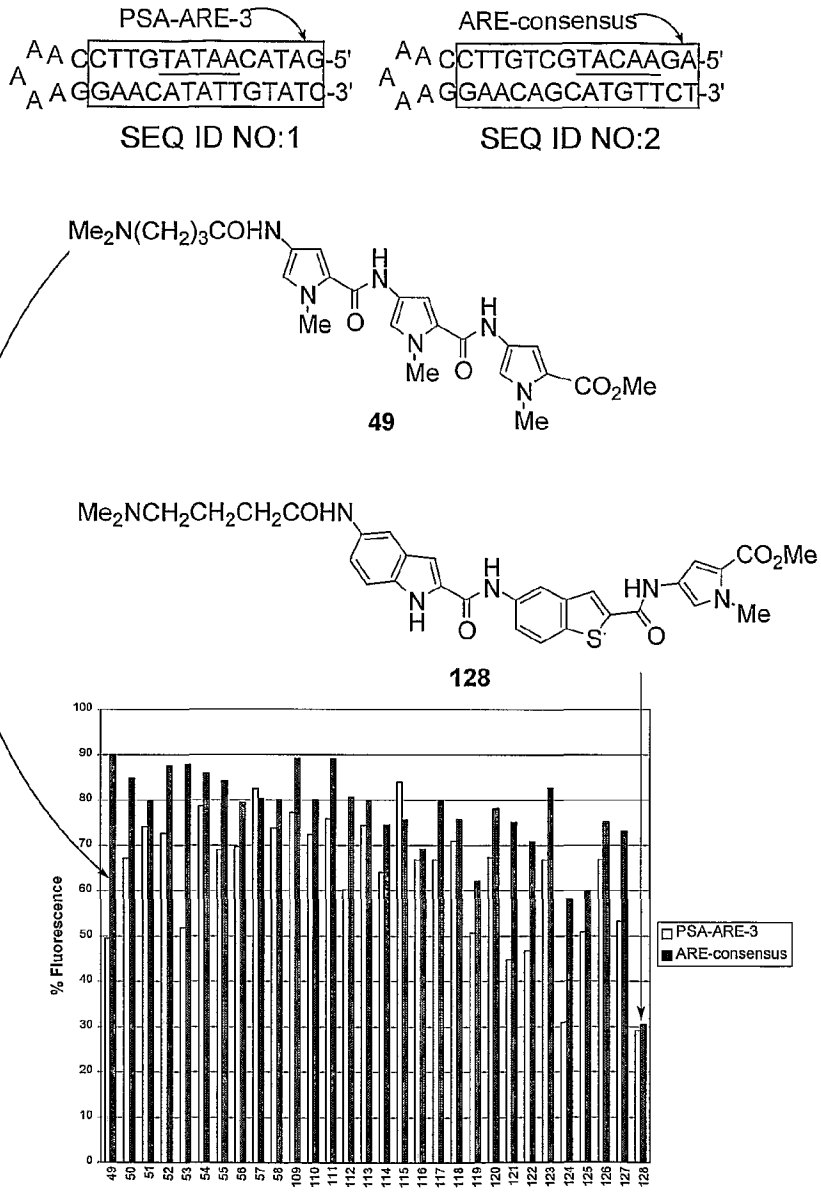

PSA-ARE-3
AA CCTTGTATAACATAG-5'
A AA GGAACATATTGTATC-3'
SEQ ID NO:1
ARE-consensus
AA CCTTGTCGTACAAGA-5'
A AA GGAACAGCATGTTCT-3'
SEQ ID NO:2
Me2N(CH2)3COHN
49
Me2NCH2CH2CH2COHN
128
% Fluorescence
PSA-ARE-3
ARE-consensus

Figure 29

% Fluorescence at 10 μM

Cmpd	PSA-ARE-3		ARE-consensus	
49	[49]	$K = 7.0 \times 10^5\ M^{-1}$	82	$K = 1.4 \times 10^5\ M^{-1}$
50	67		80	
51	74		88	
52	73		88	
53	[52]		86	
54	79		90	
55	68		87	
56	70		79	
57	82		78	
58	74		80	
109	77		89	
110	72		80	
111	76		89	
112	60		81	
113	74		80	
114	64		74	
115	84		76	
116	67		69	
117	67		80	
118	71		76	
119	51		62	
120	67		78	
121	45		75	
122	47		71	
123	67		83	
124	[31]	$K = 7.7 \times 10^5\ M^{-1}$	69	$K = 4.5 \times 10^5\ M^{-1}$
125	51		61	
126	68		75	
127	53		73	
128	[29]	$K = 3.2 \times 10^6\ M^{-1}$	[31]	$K = 4.5 \times 10^6\ M^{-1}$

Figure 30

SEQ ID NO:3

5'-CGNNNNNC A A
3'-GCNNNNNG A A A

N = A, T, G, C

5'-CG**ATGCA**C A A 3'-GC**TACGT**G A A A	≡	5'-CG**TGCAT**C A A A 3'-GC**ACGTA**G A A
A	Redundant	**B**
SEQ ID NO:4		SEQ ID NO:5

Figure 31

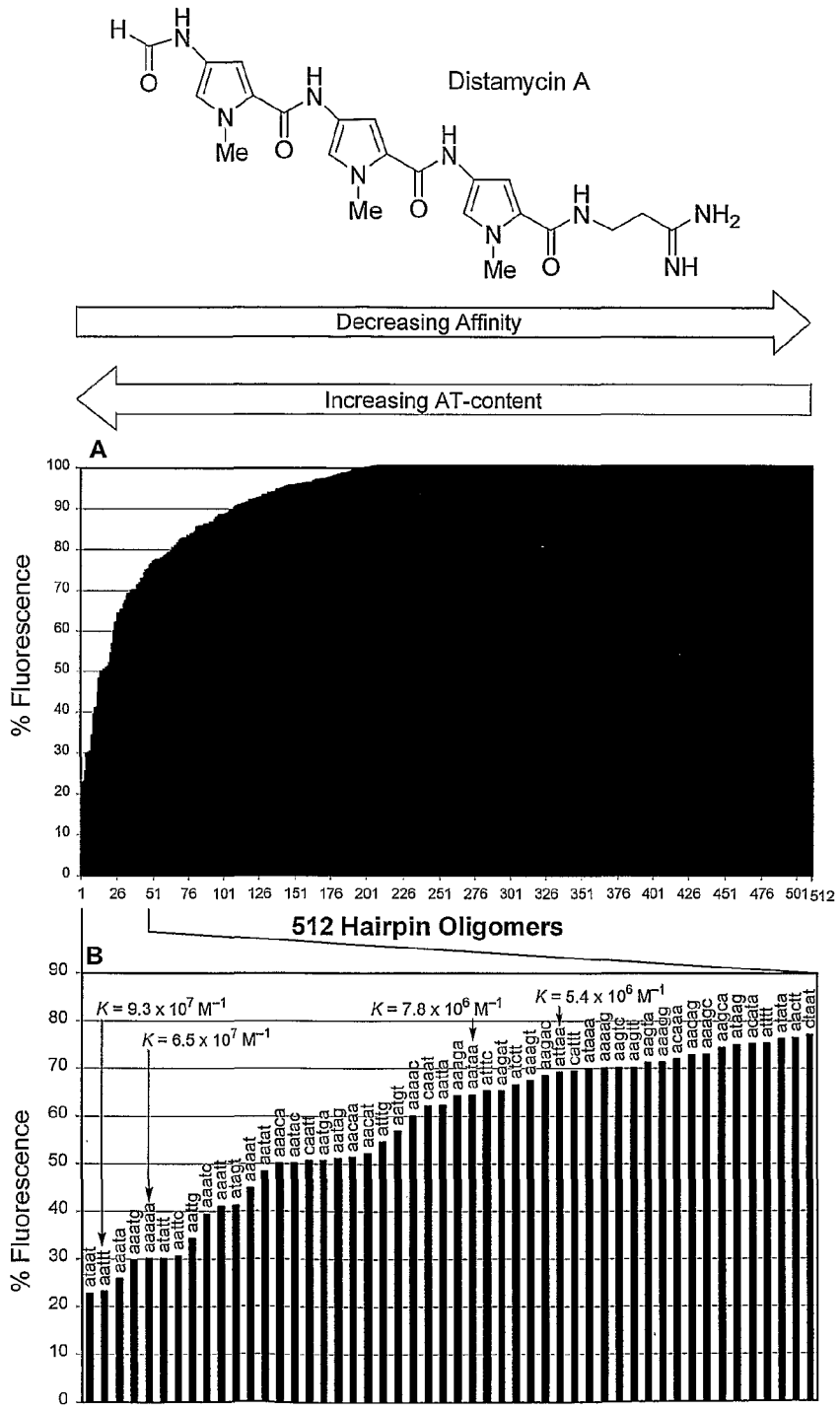

Distamycin A
Decreasing Affinity
Increasing AT-content
A
% Fluorescence
512 Hairpin Oligomers
B
% Fluorescence
K = 9.3 x 10^7 M^-1
K = 6.5 x 10^7 M^-1
K = 7.8 x 10^6 M^-1
K = 5.4 x 10^6 M^-1

Figure 32

DNA sequence	K (M^{-1})	K (M^{-1})$^{lit.}$
5'-AATTT-3'	9.4×10^7	[a] 3.1×10^7
5'-AAAAA-3'	6.5×10^7	[b] 2.6×10^7
5'-AATAA-3'	7.8×10^6	[b] 1.4×10^7
5'-ATTAA-3'	5.4×10^6	[b] 1.9×10^6

[a] Calorimetry, Rentzepris, D.; Marky, L. A.; Dwyer, T. J.; Geierstanger, B. H.; Pelton, J. G.; Wemmer, D. E. *Biochemistry* **1995**, *34*, 2937.
[b] Footprinting on a close analogue of distamycin A, Wade, W. S.; Mrksich, M.; Dervan, P. B. *Biochemistry* **1993**, *32*, 11385.

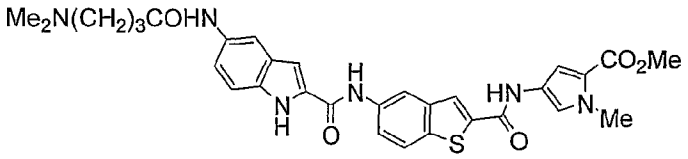
Figure 33

Ethidium bromide binding constants[a]

polynucleotide	K_{EB} (x 10^6 M^{-1})
poly[dAT]–poly[dAT]	9.5
poly[dA]–poly[dT]	0.65
poly[dGC]–poly[dGC]	9.9
poly[dG]–poly[dC]	4.5
poly[dAC]–poly[dGT]	9.8
poly[dAG]–poly[dCT]	1.3
Calf Thymus	10

[a] Baguley, B. C.; Falkenhaug, E.-M. *Nucleic Acid Res.* **1978**, *5*, 161.

Figure 34



128

ARE-consensus

5'-AGAACATGCTGTTCC A A A
3'-TCTTGTACGACAAGG A A

SEQ ID NO:2

PSA-ARE3

5'-GATACAATATGTTCC A A A
3'-CTATGTTATACAAGG A A

SEQ ID NO:1

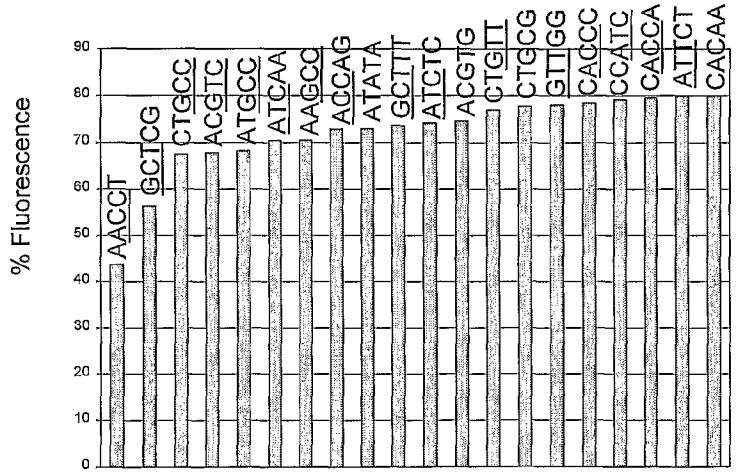
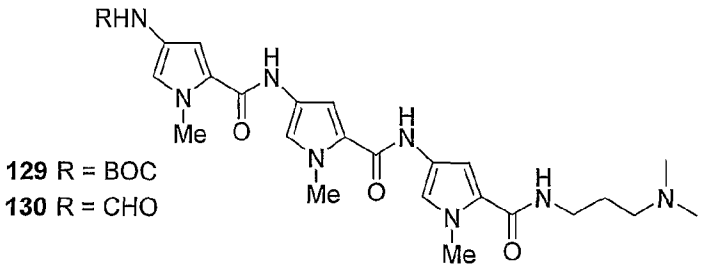
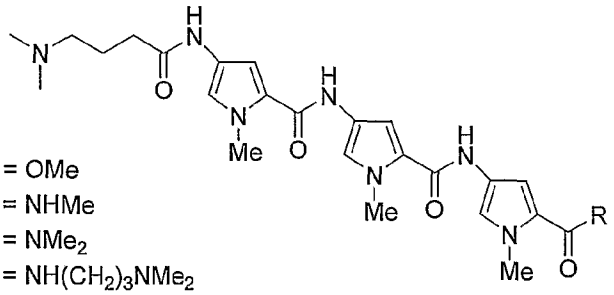


Figure 35



129 R = BOC
130 R = CHO



49 R = OMe
131 R = NHMe
132 R = NMe_2
133 R = $NH(CH_2)_3NMe_2$

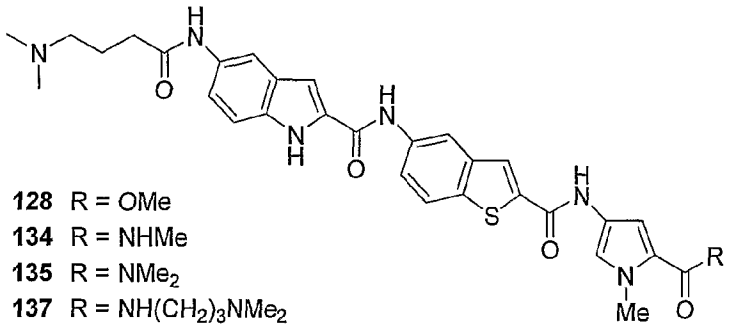
	K (x 10^6 M^{-1})[a]		ΔG[b]	
	[dA]–[dT]	[dG]–[dC]	[dA]–[dT]	IC_{50} μM (L1210)
Dist	15.0	0.071	–9.78	42
49	5.9	0.073	–9.22	44
129	2.1	0.083	–8.74	32
130	15.9	0.089	–9.80	100
131	11.8	0.110	–9.64	100
132	2.5	0.076	–8.72	100
133	12.7	0.150	–9.69	>10

[a] K calculated from $K = K_e$[EtBr]/[Agent], see ref c.

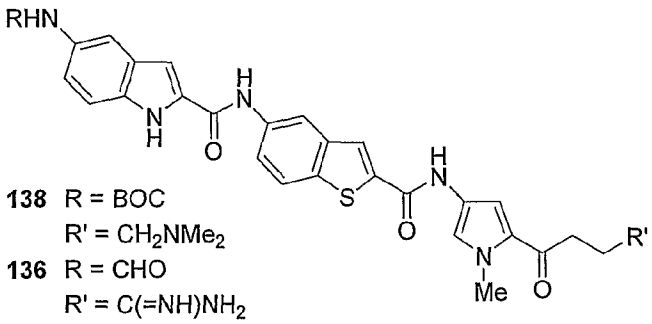
[b] $\Delta G = -RT\ln K$ (298 K)

[c] (1) *Drug-DNA Interactions Protocols*; Fox, K. R., Ed.; Methods in Molecular Biology; Humana Press: Totowa, New Jersey, 1997; Vol. 90. (2) Jenkins, T. C. Optical Absorbance and Fluorescence Techniques for Measuring DNA-Drug Interactions. In *Drug-DNA Interactions Protocols*; Fox, K. R. Ed.; Methods in Molecular Biology; Humana Press: Totowa, New Jersey, 1997; Vol. 90, p. 195.

Figure 36



128 R = OMe
134 R = NHMe
135 R = NMe_2
137 R = $NH(CH_2)_3NMe_2$



138 R = BOC
R' = CH_2NMe_2
136 R = CHO
R' = $C(=NH)NH_2$

	K (x 10^6 M^{-1})[a]		ΔG[b]	
	[dA]–[dT]	[dG]–[dC]	[dA]–[dT]	IC_{50} μM (L1210)
128	5.6	1.9	–9.20	17
138	2.5	1.3	–8.72	0.43
134	8.6	1.9	–9.45	19
135	2.5	1.2	–8.72	23
136	7.3	2.3	–9.35	0.72
137	9.5	1.5	–9.51	33

[a] K calculated from $K = K_e$[EtBr]/[Agent], see ref c.

[b] $\Delta G = -RT\ln K$ (298 K)

[c] (1) *Drug-DNA Interactions Protocols*; Fox, K. R., Ed.; Methods in Molecular Biology; Humana Press: Totowa, New Jersey, 1997; Vol. 90. (2) Jenkins, T. C. Optical Absorbance and Fluorescence Techniques for Measuring DNA-Drug Interactions. In *Drug-DNA Interactions Protocols*; Fox, K. R. Ed.; Methods in Molecular Biology; Humana Press: Totowa, New Jersey, 1997; Vol. 90, p. 195.

Figure 37

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/19404

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :Please See Extra Sheet.
US CL :548/311.4, 312.4, 433, 454; 514/397, 398, 411, 414, 415
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 548/311.4, 312.4, 433, 454; 514/397, 398, 411, 414, 415

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Please See Extra Sheet.

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

NONE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,786,377 A (GARCIA et al.) 28 July 1998, see entire document.	1-16, 19
A	US 5,401,851 A (BOYD et al.) 28 March 1995, see entire document.	1-16, 19
A	US 5,475,092 A (CHARI et al.) 12 December 1995, see entire document.	1-16, 19

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search	Date of mailing of the international search report
18 SEPTEMBER 2001	11 OCT 2001
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer FLOYD D. HIGEL Telephone No. (703) 308-1235

Form PCT/ISA/210 (second sheet) (July 1998) ★

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/19404

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (7):

C07D 233/54, 405/14, 407/14, 403/14, 413/14, 487/02; A61K 31/478, 403, 4045; A61N 35/00

B. FIELDS SEARCHED
Documentation other than minimum documentation that are included in the fields searched:

INDEX CHEMICUS
CURRENT ABSTRACTS OF CHEMISTRY
CHEMICAL ABSTRACTS